A Trend Analysis on Sri Lankan Politics based on Facebook User Reactions

M.S.Sally

2018

A Trend Analysis on Sri Lankan Politics based on Facebook User Reactions

A dissertation submitted for the Degree of Master of Computer Science

M.S.Sally

University of Colombo School of Computing

2018

DECLARATION

The thesis is my original work and has not been submitted previously for a degree at this or any other university/institute.

To the best of my knowledge it does not contain any material published or written by another person, except as acknowledged in the text.

Student Name: Ms. M.S.Sally

Registration Number: 2015/MCS/066

Index Number: 15440667

_____

Signature: Date:

This is to certify that this thesis is based on the work of Ms. M.S.Sally

under my supervision. The thesis has been prepared according to the format stipulated and is of acceptable standard.

Certified by:

Supervisor Name: Dr. M.I.E.Wickramasinghe

_____

Signature: Date:

ABSTRACT

The rapid growth of user interactions in social media sites gives useful insights on many areas. Facebook is the most popular social media site lately, with the highest number of active users, which is a valuable and hassle free source in obtaining data. Despite its enthusiastic nature, it is a mere fact that people use Facebook to gain instant updates on current news. The extreme ease of providing feedbacks using gesture based reactions is another main reason for its popularity. Politics has been a ubiquitous topic in the world at all times. Sri Lanka has been in an ethnic war for nearly three decades, followed by a nepotistic governance which lasted for nearly a decade has influenced the citizens' political conviction heavily. The "Good-Governance" trounced the previous government on such a background, which they claim to direct Sri Lanka towards a sustainable, stable, responsible and moral society with necessary constitutional amendments guaranteeing democracy to all ethnic groups eradicating corruption, wastage and fraud. The interest and motivation of this thesis builds up to discover whether there are any significant trends in the Sri Lankan political context in the perspective of the general public. The analysis of this study reveals an increasing trend in politics from 2011 to 2018. Further, it is identified that the present government has a decreasing trend over the past three years in the sight of its citizens, although they pledged for a better governance. On the contrary, the previous government has an increasing trend even though they were overpowered by the "Good-Governance" for its unscrupulous ruling.

ACKNOWLEDGEMENT

First and foremost, I would like to praise God, the Almighty for providing me the strength, knowledge and opportunity to undertake this research study and complete it satisfactorily amidst the utmost hardships through these years.

At the very outset, my sincere gratitude is accredited to my supervisor Dr. M.I.E.Wickramasinghe for his extensive support and assistance on discussing and accurately reviewing all the research work of this thesis. His immense knowledge and encouragement was a true motivation and direction to overcome all the difficulties arose throughout this year. Besides my supervisor, I would like to thank the rest of the staff of University of Colombo School of Computing including all the lecturers who imparted their valuable knowledge to us. In particular, I would like to thank our project coordinator Dr. Kasun Gunawardena for his efforts in scheduling flexible deadlines whenever it was possible and requested.

I thank my fellow batch mates who were a strength all the time. I would never fail to express my gratitude to my bosom friends who uplifted and supported me through their kind words whenever I needed at the toughest times. I would also like to thank each and every one who supported me despite their time, in filling my survey form in a short period which I know is a boring and annoying task.

Last but not the least, I would like to thank my mother for her continuous encouragement, blessings and prayers throughout my Master's Degree Programme. I wish my father was here today, to share the happiness of my achievement. His final words has been an inspiration whenever I felt bad in life. I would also like to thank my siblings, their husbands and my loving little nieces for supporting me spiritually throughout my journey so far.

Many Thanks to all of you, this thesis would not have even been possible without any person I named!

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATION

API – Application Programming Interface

CSV - Comma Separated Values

HTTP – Hypertext Transfer Protocol

JSON – JavaScript Object Notation

PHP – Hypertext PreProcessor

SDK – Software Development Kit

URL - Uniform Resource Locator

CHAPTER 1 – INTRODUCTION

Politics is a widely spoken topic in all parts of the world. It is mainly important because the critical decisions taken on the entire country solely depends on the government which is in power. As Sri Lanka consists of many diverse ethnic groups and is spread over the country it has been a prime factor to take equitable decisions that would be fairly applicable to all the citizens. In past decades, there have been many tragic and crucial incidents reported which has eventually deteriorate the development, culture and well-being of the country. The civil war that lasted for nearly three decades and the nepotistic governance followed by then, are the recent consequences caused due to unjust political decision making of the past government. Thousands of innocent civilians and the generation propagated after, were the victims due to these phenomenon.

1.1 MOTIVATION

Although Sri Lanka has been a democratic country since 1972, the direction of politics has been chaotic and a crisis more often. Some of the issues arose from the reign of the former government are, corruptive nature in the state institutions, security threats of extremist groups, rising inequality, high development debts and lack of media freedom. The present government came into power amidst such skepticism with a new revolution called the “Good-Governance” to restore democracy and rebuild the nation from where it was left. At the milestone of the third successive year of the “Good-Governance” the current political situation is yet debatable among people, as some argue it as a real political transformation and some don’t.

1.2 STATEMENT OF PROBLEM

As noted in the previous sections, the motivation for this research has been evoked in a background which people claim that the present government has failed in achieving the needs of the country and its crack down on bribery and corruption in accordance to their hundred day election manifesto of stability. This study is conducted to find whether there is any significance of the “Good-Governance” compared to the previous reign from the perspective of the general public using their interactions on social media. Social media sites have been spread rapidly for its high interactive nature and ease of use among people, despite their age, religion, gender, education level and cultural differences since its inception. Facebook, which is the most popular social media site used by billions of active online users dispersed across the globe at present, has been preferred over the other sites for this purpose.

1.3 OBJECTIVES

Facebook has played a silent but a vital role in the election campaign in favor of the present government to gain victory over the bitterly divisive decade-long ruling. It was a noticeable fact that many news and media establishments had replenished their morality by then and actively engaged in their duties of reporting the accurate information to the public. On the extreme popularity of Facebook, a news post published can reach a vast number of online users in a minimum time. Upon the ease of interacting, these posts can be viewed and given feedback as likes, comments or share. Facebook reactions, which is a novel feature introduced as an extension to the like button, is a quick, easy and gesture based way of providing feedback to a post.

As the main goal of this research is to identify the trends in Sri Lankan politics over a substantial period of time including governance of the past and present, the following objectives are intended to be performed.

- Identification of any pattern in the user involvement on politics over a period of time
- Comparison of the user engagement in political news posts over other news aspects
- Find any pattern in the freedom of expressing views in public proponent or opponent to the government
- Find the satisfaction of citizens on the governance of the present government compared to the previous by the work that is been carried out

1.4 SCOPE

This study focuses solely on news posts that are related to the political context on a Facebook news page. Data regarding user interactions such as the reaction counts, comments count and share count are obtained for each news post on a time period from 2011 to 2018 adhering to certain limitations on Facebook. As Sinhala is the native language in Sri Lanka, it was observed that Sinhalese newspapers were liked by users more over newspapers of other languages. “Lankadeepa” is chosen as the source page for data retrieval ascribed that it has a higher user interaction compared to other news pages.

1.5 THESIS OUTLINE

The rest of the thesis is organized as the follows.

Chapter 2: Describes the background study of the existing work done by previous researchers on the underlying subject area and the identified research gap.

Chapter 3: Describes the methodology, assumptions and overall approach of this research including Facebook Graph API querying, data collection and categorization into sub levels, bias analysis and calculation in a descriptive and detailed manner. It also includes a critical analysis for the news posts which compares the previous and present governance.

Chapter 4: Provides a comprehensive trend analysis for political data in order to evaluate the patterns identified in Chapter 3.

Chapter 5: Concludes the thesis by summarizing the achieved results. It also points out the open issues uncovered in this thesis which represent the intended future work.

CHAPTER 2 – LITERATURE REVIEW

Following the problem domain and the objectives of this research, it is of vital importance to identify the existing work that has been carried out by previous researchers on this area. This chapter enhances research done before and how data obtained through Facebook has be used to derive potentially valuable insights.

2.1 EXISTING WORK

Social media data has been used to obtain beneficial conclusions on many areas. Many studies highlight the importance of users in a social media site and how they generate massive data without any hassle that is favorable in making decisions. Cristian Bucur, in [1] proposes a software system design to a web platform by conducting various methods of analysis of information from web pages, which extracts opinions of customers from the online content available and support the decision making in a company. He also emphasizes how opinion mining methods would be beneficial in an era where social media is growing rapidly and contains an enormous amount of relevant information in the creation of effective referral systems, financial analysis, market research and product development.

Saurin Dave and Prof. Hiteishi Diwanji highlights the importance of social media data such as comments and reviews to study about the community [2]. The authors point out the relationship between people, how they take certain decisions based on reading the comments or reviews of others. A sentimental analysis based approach is used to find the polarity of user text. They also focus the attention on the hybrid approach of community detection in trend analysis which they regard as a lacking area in recent research conducted on this area.

The importance of microblogging and social communication in business branding is discussed by Michail Salampasis et.al in [3]. They emphasize the high popularity of usage in social media sites being a major factor to do a continuous monitoring and mining on the consumer behavior. A dataset of thousands of microblogging messages containing comments, sentiments and opinions about food and brand products has been used. The results concludes that there is a strong indication in the role of a user to support and enhance important business processes.

Despite of the business area, social media can be used as a learning tool due to its higher user interactive nature. Christopher G.Reddick et.al suggests a learning framework for local government departments [4]. The framework uses social media text analytics from Facebook pages of government departments and it integrates double-loop learning theory. In the double-loop learning theory, the first loop is used to educate the citizens through Facebook posts and

the second loop is used to get their feedbacks. The resultant insights have shown a promising direction of enhancing the quality of the public services.

Considering the above, it is a noticeable fact that a variety of sentimental analysis and data mining techniques have been used in this area of research. In a survey done by Mariam Adedoyin-Olowe et.al, they point out the challenges, issues in these techniques in areas of data size, noise and dynamism [5]. The authors proposes a method as Transaction-based Rule Change Mining (TRCM) which was used to discover the rule trend of tweets' hashtags over a consecutive period of time.

News Channels have been of one major reasons people access Facebook. A study [6] has been done by Chaker Mhamdi et.al on Fox News, CNN, and ABC news channels to find interesting relationships among the news items they post. The data were retrieved using FacePager and analysis through RapidMiner, which indicated that USA elections news have been the highest coverage among other news across all these news channels. In fact, it divulges the importance of Facebook news pages in the political background.

Political context has also been a fascinating area in social media related research lately. Many researchers find politics as a hot topic at any time and it has focused them in finding interesting relationships. In [7] Diego Tumitan and Karin Becker finds whether user comments on online newspapers reflect external indicators of public acceptance. The study is done on Brazilian politics and sentimental analysis on comments in Portuguese. The results indicated that most comments refers to frustration about politics in general; transference of opinion by the candidate's association to other corrupt politicians or parties; political scandals; poor previous administration.

In a study done by Sethunya R Joseph et.al, they explain [8] how social media has been an influencial factor in the political landscape with examples of incidents and events which took place in the world by performing several document analysis. They emphasize the unimaginable influential power that social media has created in which to affect something positively or negatively.

In addition to the above, Julia K. Woolley et.al states in their study [9] , that social media is a viable tool for political communication. The authors used Facebook group pages for this research from the beginning up to the election date that focused on how the presidential candidates Barack Obama and John McCain would be portrayed. The results of the research after a qualitative analysis concluded that Barack Obama was portrayed more positively than

John McCain which had also been the outcome in reality. Therefore, it has the proved the fact that the population could be inferred by the users engaged in Facebook pages.

In a study done by Juliet E. Carlisle and Robert C. Patton [10], the authors conceptualizes the political engagement on Facebook as focusing on the user activities regarding politics during the US presidential primary and general elections held in 2008. They use a resource model to test whether the factors helpful in understanding offline political participation also explain political participation in Facebook. In addition to that, the resources socioeconomic status [SES] and political interest has been considered to test whether network size works to increase political activity. The finding concluded that the individual political activity in Facebook is not as extensive as popular as the accounts suggest.

Stefan Stieglitz et.al [11] describes in their study, how important social media has been in communication and how it has increased the political participation by direct dialogs with the users. The statement of problem in this research is that the lack of systematic tracking and analysis approaches in the political domain with appropriate scientific methods and techniques due to the large amount of social media platforms and complex bulk information. As a solution, a methodological framework for social media analytics has been proposed in the political context.

In study done by Andrea Ceron et.al [12] on forecasting electoral results based on social media platforms, the authors state that the growing usage of social media is sharply increasing and that it could be used as a device to explore and track political preferences. They have applied a method proposed by social scientists to three different scenarios, by analyzing on one side the online popularity of Italian political leaders throughout 2011, and on the other the voting intention of French Internet users in both the 2012 presidential ballot and the subsequent legislative election. Although, Internet users are not necessarily representative of the whole population of a country's citizens, yet the analysis has shown a remarkable ability for social media to forecast electoral results, as well as a noteworthy correlation between social media and the results of traditional mass surveys. They also illustrate that the predictive ability of social media analysis strengthens as the number of citizens expressing their opinion online increases, provided that the citizens act consistently on these opinions.

In a study done by Sam H. Dekay [13] on how large companies react to negative Facebook comments, he found that, of Fortune 100 companies 79 percent use some form of social media to communicate with customers and other stakeholders and that most marketing specialists recommend that negative comments should be treated as opportunities to resolve potential

problems. The study identified official Facebook pages sponsored by the top ten companies (as defined by the Forbes 2000 for 2010) in four industry groups – banking, retailing, software and services, and household and personal products. The number of negative comments posted to these pages is calculated and corporate reactions to the comments analyzed. The study concludes that large corporations do not generally approach negative comments as public relations opportunities, but prefer to censor, or ignore, critical feedback.

Tamara A. Small [14] has performed a study to analyze on Canadian politics using Twitter Hashtags. The author has selected Twitter as the source of data as it was the latest social networking tool since then, and is said to be reshaping politics. This research analyzes the intersection of microblogging and Canadian politics and seeks to fill this gap through a content analysis of the most popular Canadian political hashtag, #cdnpoli. With 50 million tweets per day, hashtags are central to organizing information on Twitter. Hashtags organize discussion around specific topics or events. The popular literature suggests that Twitter is a democratic media because it allows for on-the-ground reporting of breaking news and democratic activism. This analysis shows that informing is the primary function of a political hashtag such as #cdnpoli.

Considering the aforementioned studies on social media analytics, it is a notable fact that there has been no analysis done lately on Sri Lankan politics to identify the trends by means of comparing the former with its latter government bodies. It is also noticeable that, the use of Facebook reactions is a lacking area that many have failed to address or unconsidered. The motivation of this research grows in a background as such.

CHAPTER 3 – METHODOLOGY AND ANALYSIS

This chapter discusses the methodology used in initial data collection from Facebook, the description of data fields, data preparation process, assumptions made, surveys conducted and their analysis and bias calculations. In addition to that, it also includes the analysis of the news items based on Facebook reactions followed by a critical discussion.

3.1 DATA COLLECTION

Data is collected from Facebook using the Facebook Graph API Explorer version 2.11. Graph API is a low-level HTTP-based API that can be used to programmatically query data. It is a representation of social graph composed of nodes, edges and fields.

Nodes – the primary elements that are found in Facebook such as a user, photo, page, post

Edges – the connection between the primary elements such as the comments on a photo

Fields - attributes of the primary elements such as the name of a user

Since the API is HTTP-based, it can be used with any language that has an HTTP library. It can be executed directly on a browser by browsing for Facebook developers URL. A GET request can be used to obtain necessary information about a node.

Figure 3.1 illustrates how the nodes are being interconnected with each other and how their fields can be accessed by the edge relationships.

Figure 3.1: Model of Facebook Graph API

The results obtained through a GET request is in JSON format and generally, it will not be received through a single response. Hence, the responses are paginated by default.

3.2 FACEBOOK GRAPH API EXPLORER

A Facebook account is essential to use Facebook Graph API. The URL https://developers.facebook.com/tools/explorer is used to access the Graph API interface. In Figure 3.2, at the initial Graph API Explorer interface the URL of the Facebook page is inserted as a GET request which will give basic details of the page such as the name and the id. The id of the page is most important in order to access the posts related to the page.

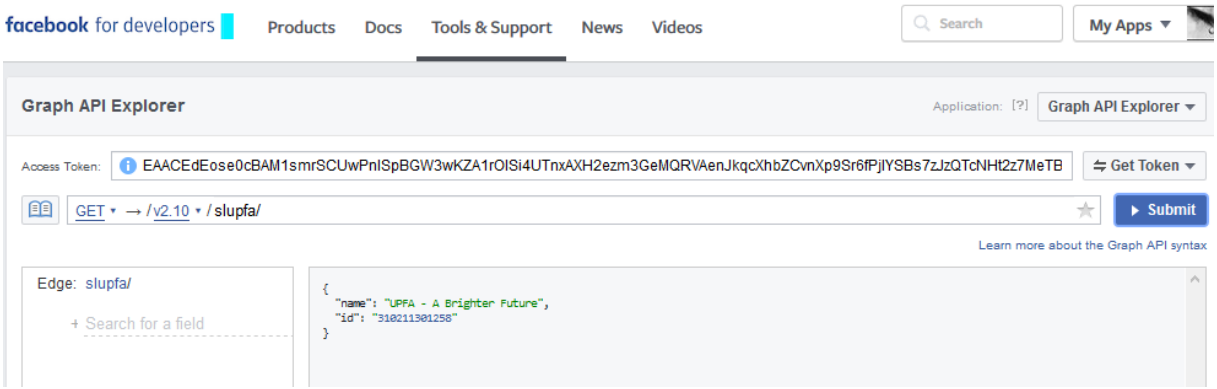

Figure 3.2: Initial screen of the Graph API Explorer

The page id can be used for further querying the nodes and edges attached to the page. As shown in Figure 3.3 it is a much concise way than using the URL itself.

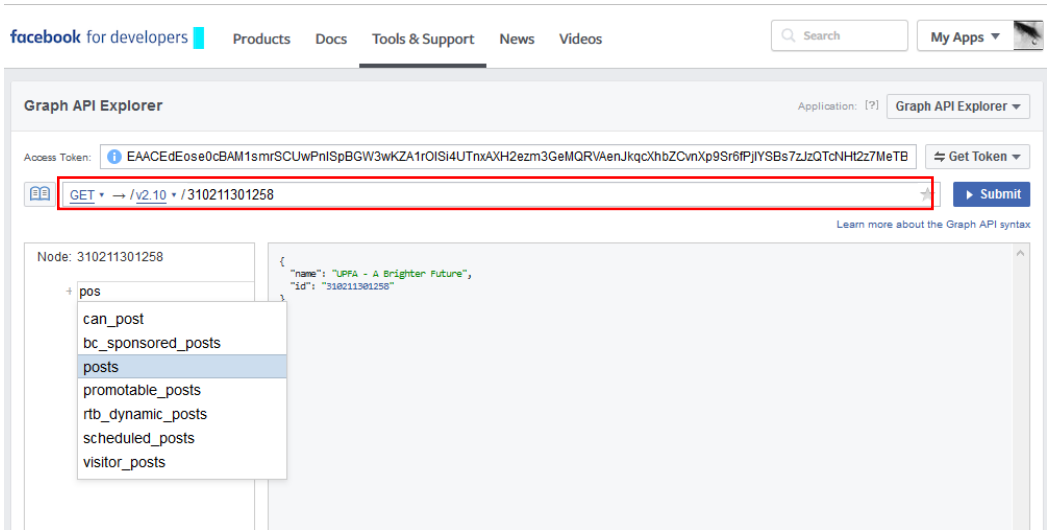

Figure 3.3: Fields attached to a page node

The attributes of the page can be accessed after the selection of the node. The interface provides a feature to search for fields of a node. Figure 3.3 shows some of the fields that are attached to the page node. The node that is considered to retain data for this research is the “post” node. A limit and an offset can be set for the node “posts”. The limit defines the number of results that can be obtained at a single request. As Graph API has a limit of hundred data items per GET

request, data will be collected on a batch basis by setting the offset. The offset is a numerical value that is set to indicate the start of each page. Offset based pagination does not concern on the chronology of the events and some results may not be returned due to certain conditions set on Facebook. The alternative method of accessing all posts is to set a time-based paginating using the “since” and “until” parameters. Regrettably, this method has been removed by Facebook since July, 2017.

In the results view, the data obtained is expected to include the fields post id, title, count of reactions on each type, time of creation and number of comments and shares.

3.3 FACEBOOK SDK FOR PHP

The results acquired from the web interface of the Graph API Explorer always forces to be 'paged' by returning the first few results, and a link to the next few. The tediousness of this task could be overcome by using a Facebook SDK which always returns the entire result set in one time. The Facebook SDK for PHP which is a library with powerful features that enable PHP developers to easily make requests to the Graph API with the integration of Facebook login, has been used in this study. A single PHP script is used to retrieve the required data as stated in the previous section and to transform it to JSON format.

3.4 PARSE DATA USING OPEN REFINE

OpenRefine is an open source tool used to handle large, messy data. The JSON data obtained is parsed using OpenRefine and is exported as an Excel sheet. Whilst parsing the data, unescaped double quotes in JSON description fields were found, and were eliminated.

The parsed data can be exported to .xls format. Figure 3.4 shows the view of the data after exported into Excel Sheet and Table 3.1 contains its description.

	_ - id	_ - title	_ - sad	_ - wow	_ - like	_ - love	_ - haha	_ - angry	_ - shares	_ - comments	_ - created_time
2	1812700011941164_1561100763958075	සූලි සිට ආසන්න පට් - වායු බැලුන නැති වාහන ආනයනය තහනම්	0	0	10	1	0	0	0	0	2018-01-19T13:15:08+000
3	1812700011941164_1561024740632344	හඳුන්නෙත්ති විශේෂ අධිකරණයක් ඉල්ලයි	0	0	20	0	0	0	3	2	2018-01-19T11:39:08+000
4	1812700011941164_1561001353968016	පොහොට්ටුවේ පෙත්සම් 14 කල් යයි	0	0	15	0	0	0	1	0	2018-01-19T11:05:09+000
5	1812700011941164_1560909833977168	මල සිරුරත් විකුණන් කෑ දෙදෙනකුට වසර 8 ක සිර දඬුවම්	0	0	24	0	0	0	0	1	2018-01-19T08:26:07+000
6	1812700011941164_1560893303978821	නාමයෝජනා පෙත්සම් විභාගය 24ට කල් යයි	0	0	4	0	0	0	0	0	2018-01-19T07:54:04+000
7	1812700011941164_1560858610648957	පාසල් සිසුන්ගේ වැඩමුළුවට ආ අපේක්ෂකයා ඉවතට	0	0	13	0	2	0	0	0	2018-01-19T06:48:12+000
8	1812700011941164_1560858583982293	මාලක සිල්වා අද උසාවියේ	0	0	12	0	0	0	0	0	2018-01-19T06:48:06+000
9	1812700011941164_1560841393984012	සම්භාවනා මුවාවෙන් කරගෙන ගිය ගණිකා මඩ්‍යස්ථානයක් එවල	0	0	8	0	1	0	0	0	2018-01-19T06:17:07+000
10	1812700011941164_1560819053986246	පොදුජන පෙරමුණේ පෙත්සම නිෂ්ප්‍රභා වේ	0	0	10	0	2	1	0	3	2018-01-19T05:44:08+000
11	1812700011941164_1560788637322621	පර්පචුවේල් ට්‍රෙෂරීස් සමාගමේ රුපියල් බිලියන 12ක මුදල් තමා ඇ	0	0	176	4	56	13	44	61	2018-01-19T04:45:10+000
12	1812700011941164_1560762450658573	කොටහේන අලුත් මාවතේ මාර්ග කිහිපයක් දින තුනක් වැසේ	0	0	6	0	0	0	0	0	2018-01-19T04:07:07+000
13	1812700011941164_1560280100706808	මම ගෙදර යන්නේ දූෂිත දේශපාලනඥයින් අවිවී මහා නරකාදියට	0	3	71	4	45	1	10	49	2018-01-18T15:41:07+000
14	1812700011941164_1560208290713989	විදුලි සේවක වර්ජනය අවසන්	0	0	22	0	1	0	2	0	2018-01-18T14:03:16+000
15	1812700011941164_1560208224047329	නිවපු නිතිසතිව සොලිසිටර් ජනරාල්ට ලෙකෝ ලේකමීට නඩු	0	0	15	0	0	0	2	1	2018-01-18T14:03:10+000
16	1812700011941164_1560181284050023	වැලිඅමුණ වාර්තාව ක්‍රියාත්මක කළා - ශ්‍රී ලන්කන්	0	0	21	0	0	0	1	0	2018-01-18T13:30:10+000
17	1812700011941164_1560118040723014	උදයංගට මිලියන 94 ක වත්කම්	0	0	58	1	2	0	8	6	2018-01-18T12:26:10+000
18	1812700011941164_1560076497393835	වැදගත්ම කොටස පාර්ලිමේන්තුවෙන් හංගලා - කැෆේ	0	0	17	0	2	1	7	2	2018-01-18T11:23:10+000
19	1812700011941164_1560055877395897	එජාප ජාතික ලැයිස්තු මන්ත්‍රී ඉල්ලා අස්වෙයි	0	0	36	0	1	0	0	1	2018-01-18T10:50:11+000
20	1812700011941164_1560055844062567	ලිංගික අල්ලස් ඉල්ලූ සැකකරු අත්අඩංගුවට	0	0	31	1	0	0	4	0	2018-01-18T10:50:07+000
21	1812700011941164_1560017844066367	ලංවිම සේවක විරෝධතා පාගමන අගමැති කාර්යාලය අසලට	0	0	13	0	0	0	2	0	2018-01-18T09:45:06+000
22	1812700011941164_1559984210736397	මාලක සිල්වා අල්ලන්න වරෙන්තු	1	0	52	1	3	0	6	6	2018-01-18T08:41:13+000

Figure 3.4: Data exported into Excel Sheet

Data	Description
_ - id	Unique identifier of a page post. The value has two components concatenated with "_". The first part is the page id and the latter is the post id.
_ - title	The title of the post. It will be in Sinhala language.
_ - sad	The six reactions with its counts.
_ - wow	
_ - like	
_ - love	
_ - haha	
_ - angry	
_ - shares	The number of occurrences the post was shared.
_ - comments	The number of comments on the post.
_ - created_time	The date and time which the post was published in Facebook.

Table 3.1: Description of data

3.5 DATA PREPARATION

It was observed that not every post was a news item but a profile or cover photo update which required elimination from the data set. The news items are subjected for categorization into many sub levels merely on my personal opinion. The main issue found here is the ambiguous nature of the news titles. In many occurrences the meaning was not expressed explicitly, due to the usage of the language and the usage of cartoons, hence more exploration was required on it.

At the higher level, the news items have been identified as a nominal variable with two values as political and the non-political. They are indicated as "1" and "0" respectively, for the ease of analyzing. The politically sorted items are then grouped depending on multiple sub levels depending on its nature which is mentioned in the latter part of this chapter. The field created time mentioned in Table 3.1, is of ISO 8601 format which requires to be transformed to a date format, in order to be accessed in a yearly and monthly manner.

3.6 ASSUMPTIONS

Categorizing news items as political has been a laborious task due to several reasons. The main reason is, it is hard to give a clear-cut definition on which news really falls in to either group or

its sub levels. It depends on my viewpoint of categorizing things which creates a bias. The other reason is, the content of some news items do not denote what is stated in its title which subsequently indicates an entirely different meaning. Hence, these vague material had to be examined separately by reading its content on Lankadeepa web portal.

In order to estimate the degree of the bias introduced, a survey has been conducted by selecting a sample based on demographics. Factors such as gender, age, religious belief and highest educational qualification were considered. The sample set is chosen to closely relate on the dispersion of religion and gender in Sri Lanka. According to 2012 census [15], the Sri Lankan population has a composition of religion as shown in Table 3.2.

Religion	Percentage of Persons
Buddhism	70.2
Hindu	12.6
Islam	9.7
Christian	7.4
Other	0.1

Table 3.2: Population by Religion

While selecting the sample, Hindus were not included due to their illiteracy in Sinhala which causes difficulties in reading and understanding the news items. Other religious beliefs considered here includes atheism, agnostism and univerlism.

The male to female ratio under each age range is shown in Table 3.3 [16]. The age range chosen for this survey is 21-40.

Age Range	Gender Ratio (Male/Female)
Total	0.96
At birth	1.04
Under 15	1.04
15–64 years	0.96

Table 3.3: Population by Gender

3.6.1 SAMPLING AND SURVEY

The sample is formed using stratified sampling method on factors of religion and gender. The strata are based on Table 3.2 excluding the Hindus and Table 3.3 considering the only

age group as 15-64 years. The selected sample was given a survey form to be filled through Facebook. Three types of survey forms were created for this purpose based on my personal viewpoint of grouping.

1. Type 1 – contains all political news items
2. Type 2 – contains all non-political news items
3. Type 3 – contains 53:47 ratio of political to non-political news items. This ratio is defined considering the total political: non-political news items after categorization.

3.6.2 ANALYSIS OF SURVEY

The graphs and tabulations referred for this analysis are created using the SPSS statistical software. The sample chosen contains a total of 47 persons consisting a mixture of Buddhists, Christians, Muslims and Other religious Beliefs. Hindus are not considered due their illiteracy of Sinhalese. Other religious beliefs include atheism, agnosticism or universalism. The statistical figures used for this analysis to formulate the sample is based on the categorical variables religion and gender is represented in Table 3.4 as a cross tabulation with its percentages on frequencies. The total percentage of Buddhists, Christians, Muslims and Other Religious Beliefs are 55.3%, 14.9%, 17%, 12.8% respectively. The total percentages of females and males are 61.7% and 38.3%. Moreover, the subsequent compositions of gender within each religious group can be viewed as well.

Religion * Gender Crosstabulation

			Gender		
			Female	Male	Total
Religion	Buddhism	Count	18	8	26
		% within Religion	69.2%	30.8%	100.0%
		% within Gender	62.1%	44.4%	55.3%
		% of Total	38.3%	17.0%	55.3%
	Christianity	Count	5	2	7
		% within Religion	71.4%	28.6%	100.0%
		% within Gender	17.2%	11.1%	14.9%
		% of Total	10.6%	4.3%	14.9%
	Islam	Count	5	3	8
		% within Religion	62.5%	37.5%	100.0%
		% within Gender	17.2%	16.7%	17.0%
		% of Total	10.6%	6.4%	17.0%
	Other	Count	1	5	6
		% within Religion	16.7%	83.3%	100.0%
		% within Gender	3.4%	27.8%	12.8%
		% of Total	2.1%	10.6%	12.8%
Total		Count	29	18	47
		% within Religion	61.7%	38.3%	100.0%
		% within Gender	100.0%	100.0%	100.0%
		% of Total	61.7%	38.3%	100.0%

Table 3.4: Cross Tabulation of the sample by Religion and Gender

The survey is conducted by providing a spreadsheet to the participants which consists a form for personal details and a list of news headlines drawn from Lankadeepa for period 2011-2018. The respondents had to mark these news items as political or non-political depending on their personal opinion. The news items chosen for the survey forms were based on the proportion relating to its year and type that were already categorized on my personal viewpoint. Consequently, the three types of survey forms created, namely form 1, form 2 and form 3 which depicts all political news, all non-political news and a combination of political and non-political news based on the ratio obtained on grouping. The ratio applied here is 55:45 consisting of political and non-political news items. All three survey forms contains a total of hundred news items. The total count of news items marked as political by each participant is considered as the observed response. The count of political news items in each survey form is the expected response. Therefore, the expected response has three values as hundred, zero and fifty five. The bias is defined as the difference between the expected and observed response. The mean bias of each survey form type and the total bias is illustrated in Table 3.5 below.

Survey Form Category	Mean Bias
All Political	52
All Non-Politcal	12
Political and Non-Political	20
Total	28

Table 3.5: Mean Bias of the survey forms

A higher bias is shown in the all political group which contains 100% political news. The bias of the other two categories are comparatively less and the total bias of the survey is 28, which indicates the average amount that deviates from my opinion. The mean difference which is the gap between the two data points of the observed and expected responses of each survey form type is shown in Figure 3.5. The visual illustration of the gaps are numerically equal to the values obtained in Table 3.5.

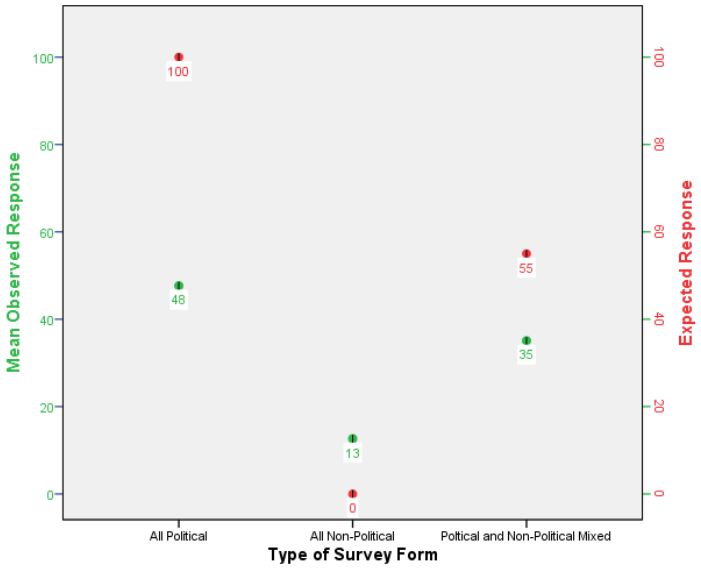

Figure 3.5: Mean Observed and Expected Response by Type of Survey Form

The personal details collected from the participants includes demographic factors such as age, gender, religion and educational background. The definitions, conditions and grouping applied on each factor is as follows,

Age - The age range for this survey is twenty one to forty with four equal ranged age groups such as,

- 21-25
- 26-30
- 31-35
- 36-40

Gender – A dichotomous variable with values male and female

Religion – The main religions in Sri Lanka including Buddhism, Christianity, Islam and Other religious belief including atheism, agnosticism or universalism. Hinduism is excluded as mentioned before.

Educational Background - The educational attainment of the participants. Three levels are been identified as follows:

- No Bachelors' Degree – This contains either Advanced Level, Diploma, Higher Diploma or any Professional qualifications
- Bachelors' Degree – This includes only Bachelors' Degree qualification
- Bachelors' Degree and Above – This includes a Bachelors' Degree and additionally following or completed a Masters' Degree, Postgraduate Diploma, Mphil or PhD.

The goal of this analysis is to find the factors that affect the political opinions of people considering my opinion as the ground truth. The mean absolute difference of responses is plotted against each factor or factors. The mean absolute difference of responses is defined as,

$$\text{Mean Absolute Difference of Responses} = \frac{|\text{Expected Response} - \text{Actual Response}|}{\text{Number of Repondents}}$$

3.6.2.1 GENDER

The graph in Figure 3.6, shows the deviation of responses by gender. It appears that females have a higher deviation than males which eventually suggest that the opinion of males are converging towards the expected response. Perhaps the higher awareness of the social consciousness males persists than females may be a fair justification for the consequence.

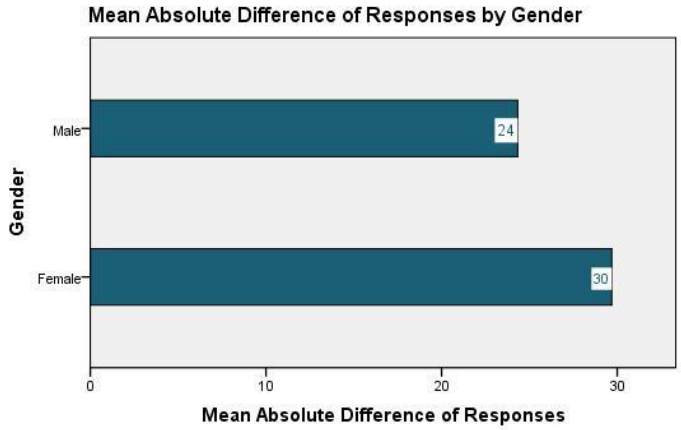

Figure 3.6: Mean Absolute Difference of Responses by Gender

The gender can be classified further by educational background as in Figure 3.7. It is visible that, males holding a Bachelors' Degree and above has the least divergence and females having only a Bachelors' Degree has the highest divergence. This can be related to the reason suggested afore regarding Figure 3.6 and additionally, the intellectuality skills that improve with the accomplishment of higher education. Thus, higher the educational attainment is, the deviation lowers.

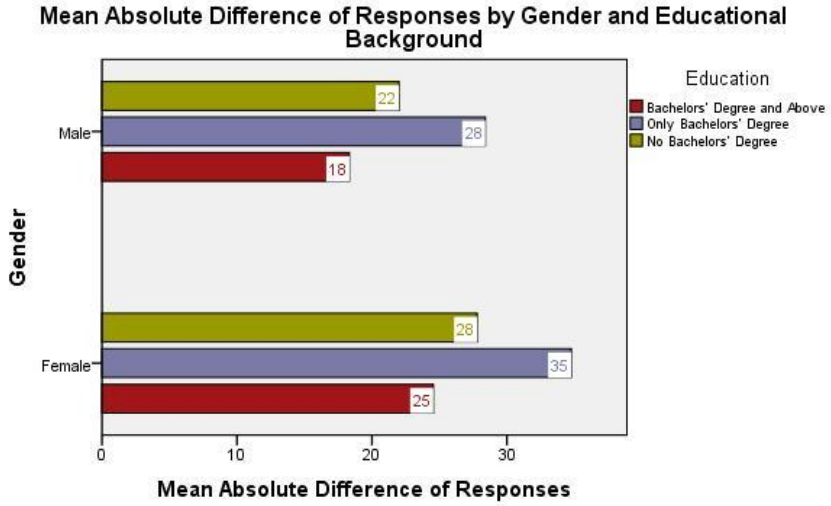

Figure 3.7: Mean Absolute Difference of Responses by Gender and Educational Background

Moreover, gender can be grouped by religion as shown in Figure 3.8. In all religions, it is seen that the male divergence is lower than the female. The highest divergence is visible in Muslim women and the lowest divergence is visible in Christian males. This may occur since Sri Lankan Muslim women intrinsically deny to acquire higher education as of certain norms set by its culture. The deviation differences between the genders within each religion is lower in Buddhists and Other religious beliefs and is higher in Muslims.

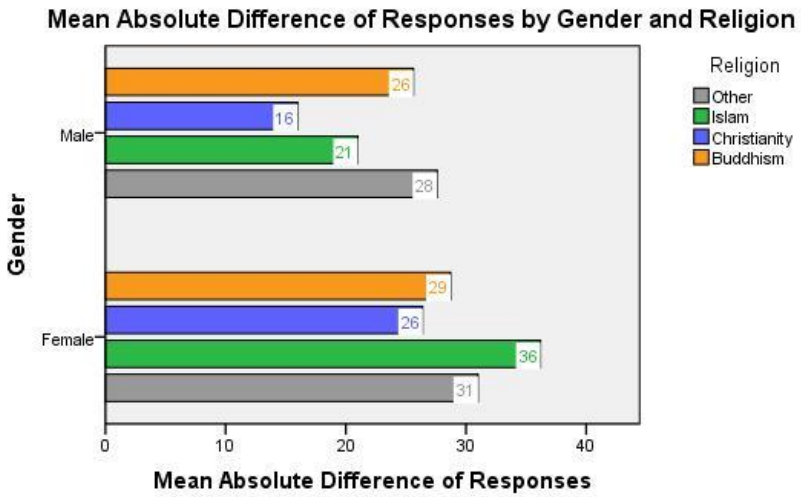

Figure 3.8: Mean Absolute Difference of Responses by Gender and Religion

3.6.2.2 RELIGION

Another factor to be considered is the religion of the participants. As in Figure 3.9, Islam shows the highest and Christianity shows the lowest divergence. In fact, the high divergence is affected

by Muslim women as seen in Figure 3.10 and the reason can be explained as noted above. The deviation of Buddhists is on par with the followers of other religious beliefs. According to the background of the participants, the followers of other religious beliefs were almost Buddhists and this may have influenced the consequence.

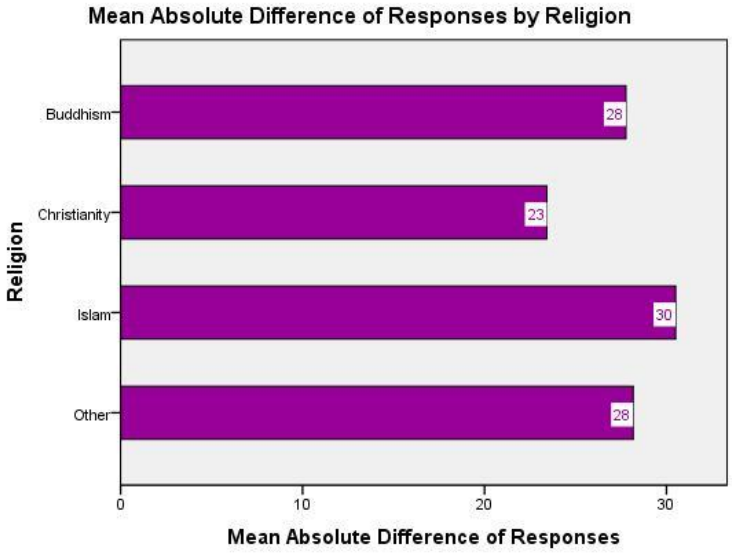

Figure 3.9: Mean Absolute Difference of Responses by Religion

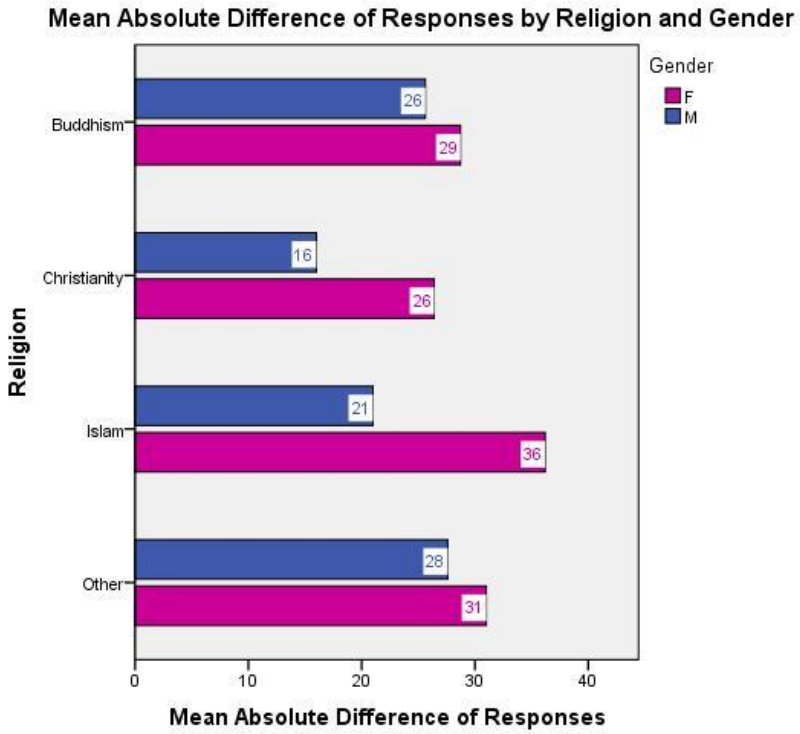

Figure 3.10: Mean Absolute Difference of Responses by Religion and Gender

Furthermore, the mean absolute difference of responses is clustered by religion and educational background as shown in Figure 3.11. In Buddhism, the deviation is higher in those who holds only a Bachelors' Degree, whereas in other religious beliefs the deviation is higher in those who does not hold a Bachelors' Degree. Muslims and followers of other religious beliefs shows a decreasing pattern among the educational levels. A lower deviation is visible among those respondents having a Bachelors' Degree and above in all religious backgrounds and Muslims hold the least deviation. A noticeable variance do not possess among the educational levels of Christians. However, Buddhists show an unusual behavior among the educational levels. A closely related reason may be the fact of the higher participation of Buddhist females in the survey.

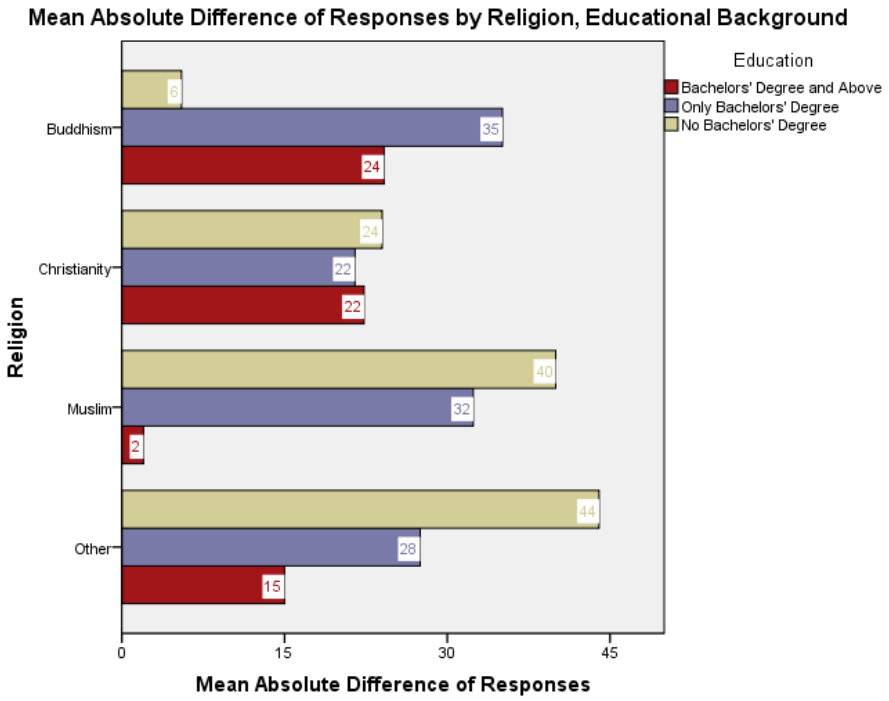

Figure 3.11: Mean Absolute Difference of Responses by Religion and Educational Background

3.6.2.3 EDUCATIONAL BACKGROUND

The educational background is a prime and interesting demographic factor for this analysis. The graph plotted in Figure 3.12 states that Bachelors' Degree and above has the lowest and only Bachelors' Degree has the highest deviation. The intellectual and reasoning skills a person develops with the attainment of higher education is a definite reason that can be given with regard to this observation. Moreover, those not possessing a Bachelors' Degree shows a lower deviation compared to those who hold only a Bachelors' Degree. A clear-cut opinion cannot be

given for this, however it can be surmised that a person drops in natural reasoning skills due to the parroting strategy followed by many Sri Lankan undergraduates.

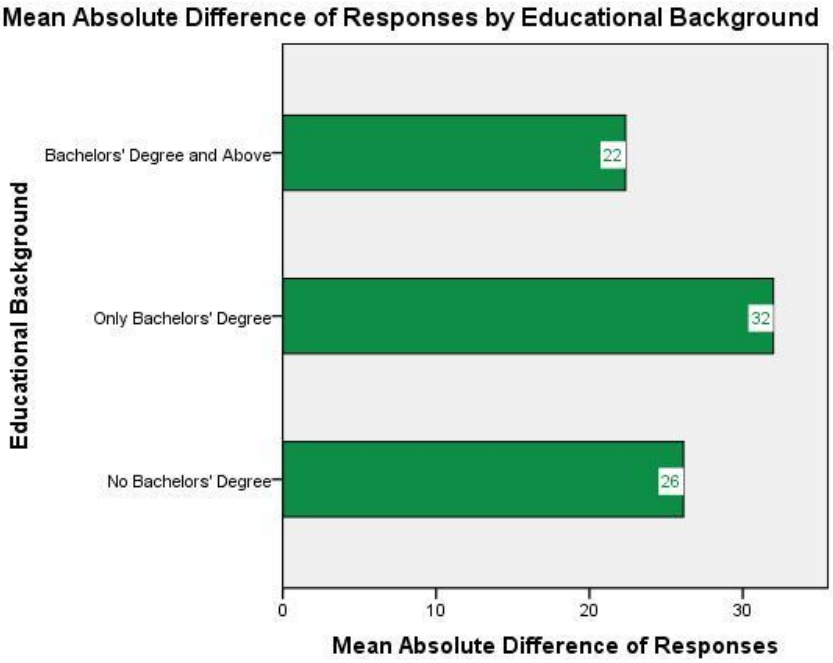

Figure 3.12: Mean Absolute Difference of Responses by Educational Background

The above results can be analyzed further with respect to gender. In Figure 3.13, the mean absolute difference of the observed and expected responses has been clustered by educational background and gender. The overall deviation is less in respondents that has a Bachelors' Degree and above. All three educational levels shows that the divergence is more in females than males. The reason to this can be formulated on the fact of social unawareness of females though having a higher educational background. Further, it strengthens this fact if the deviation of females having Bachelors' Degree and above are compared with males having only the Bachelors' Degree, which could be explained as women can be smarter than men of a certain level of education by acquiring more knowledge to combat their social unawareness. Apart from their educational level, their marital status may affect as well. As Sri Lankan women they fully engage and devote themselves in fulfilling their duties towards their family, hence find hardly any time to acquire further knowledge.

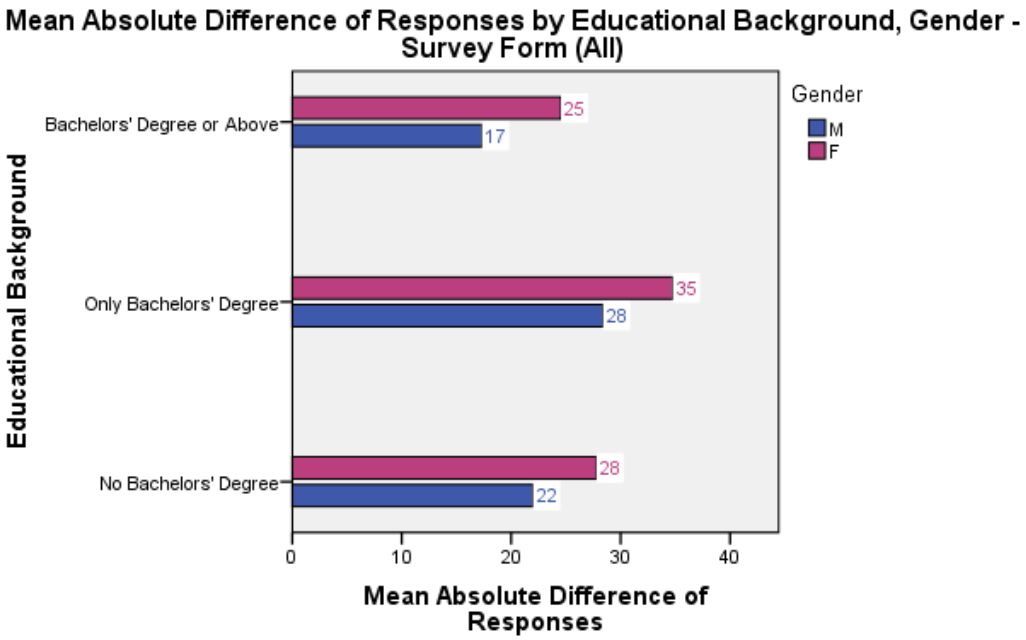

Figure 3.13: Mean Absolute Difference of Responses by Educational Background and Gender

Moreover, the educational background can be classified by religious aspect as in Figure 3.14. Buddhists not having a Bachelors' Degree shows low divergence compared to other levels of education. In contrary, Christians, Muslims and followers of other religious beliefs has a pattern that declines with higher attainment of education. However, Muslims with Bachelors' Degree and above show the least deviation among all.

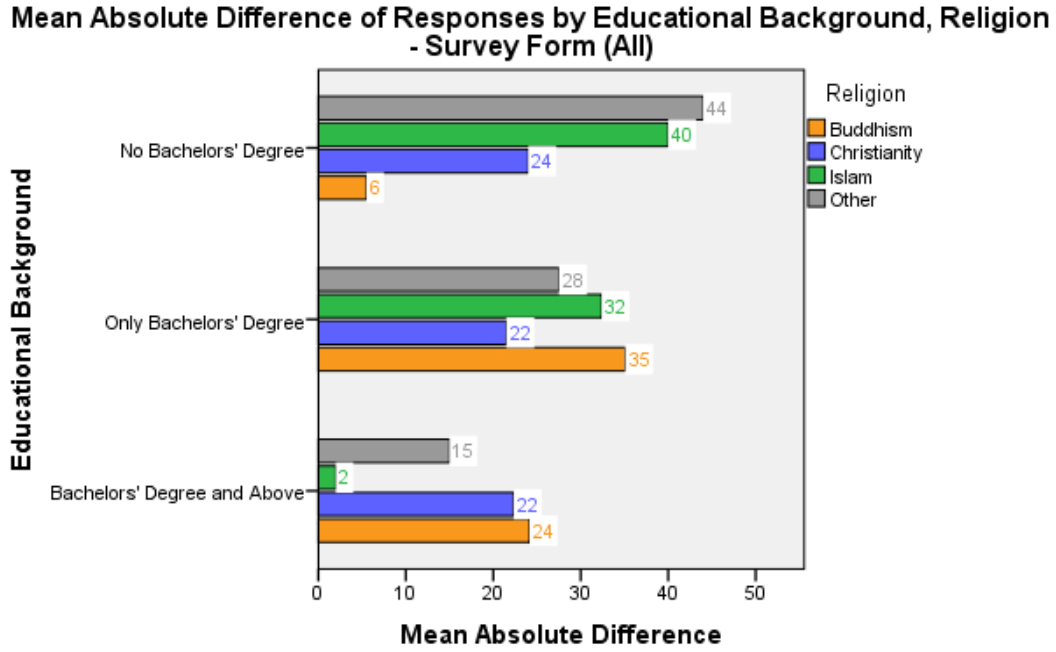

Figure 3.14: Mean Absolute Difference of Responses by Educational Background and Religion

3.6.2.4 AGE GROUP

The mean absolute difference of responses by age group illustrated in Figure 3.15, shows that the highest divergence is in age group 21-25 and the lowest divergence is in age group 31-35. This can be since many individuals of 31-35 are engaged or inclined on higher education whereas individuals of 21-25 are lack in intellectual skills and are inexperienced. Age group of 36-40 has a higher deviation compared to 31-35 and this may be due to the cognitive aging. Categorizing it further over gender as in Figure 3.16, males in age group 31-35 shows the least deviation and males in age group 36-40 shows the highest deviation. The same behavior is visible in females over all age groups.

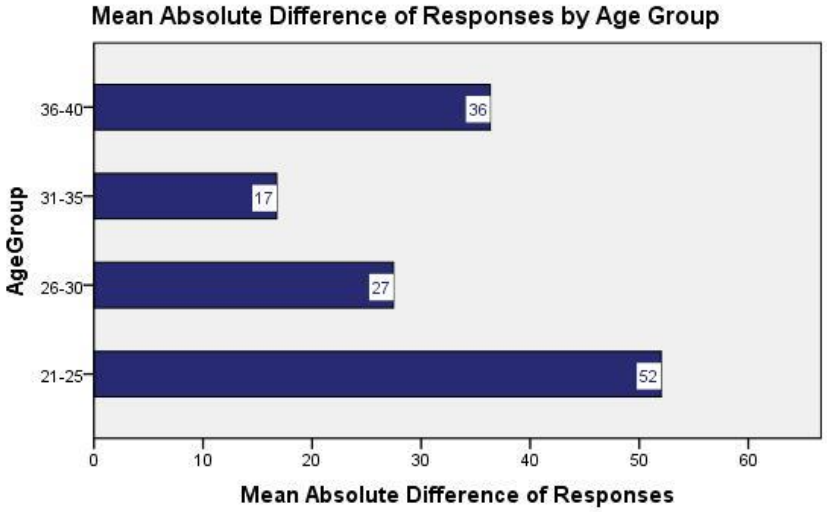

Figure 3.15: Mean Absolute Difference of Responses by Age Group

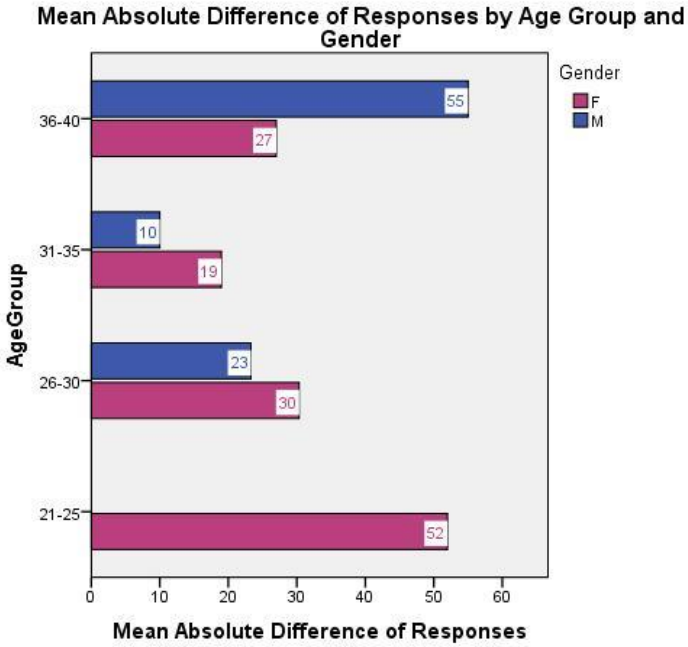

Figure 3.16: Mean Absolute Difference of Responses by Age Group and Gender

Furthermore, the deviation can be categorized by age group and educational background as illustrated in Figure 3.17. It manifests that the least divergence occur in the age group 31-35 having a Bachelors' Degree and above and the highest divergence occur in the age group 21-25 with no Bachelors' Degree. A noticeable pattern is not observed in the respondents which do not have a Bachelors' Degree. However, those who hold only a Bachelors' Degree or above has a significant pattern of decrement.

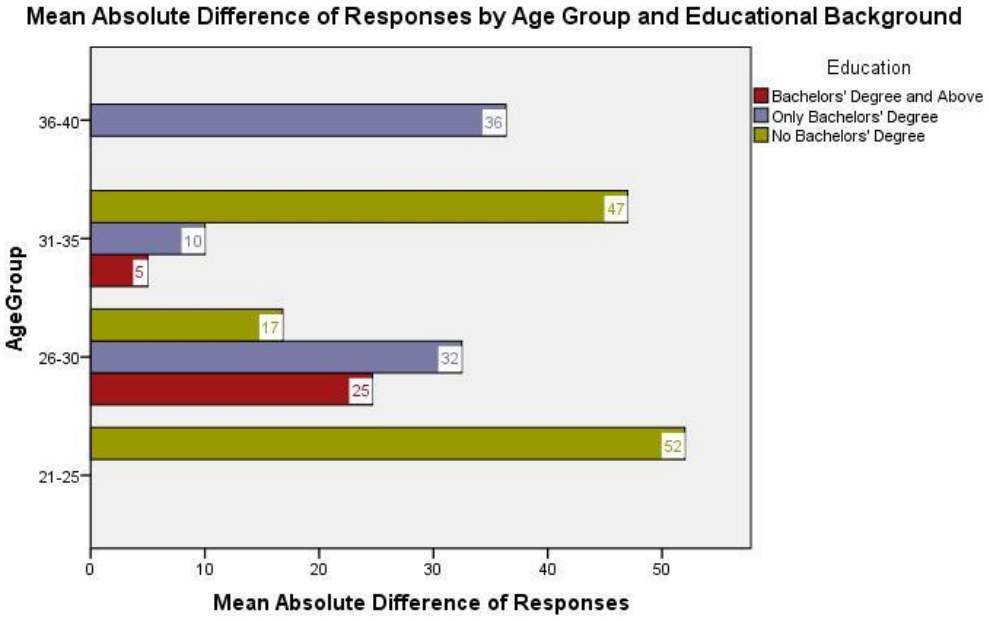

Figure 3.17: Mean Absolute Difference of Responses by Age Group and Educational Background

The survey analysis summarizes highlighting the facts that the educational background, age and gender are prime determinants that affect the cognitive reasoning ability of a person despite the religion. Certain norms set by religious groups can be an influential factor in acquiring knowledge. Hence, religion can be regarded as indirectly affecting the intellectuality of a person.

3.6.3 ANALYSIS OF THE NEWS POSTS

The total news posts gathered for this research is 4626 from August, 2011 to March, 2018. The data contains news items reported at the periods of two governments, namely the past and the present government. The entire duration from 2011 to 2018 is subdivided into two segments considering the period of ruling of the past government (2011-2015) and the present government (2015-2018). The news posts are mainly categorized as political and non-political and the composition is illustrated in Table 3.6 with 46.3% of non-political and 53.7% of

political news items. Significantly, the political news items reported within the time frame considered is higher compared to non-political news items.

Statistics of News Items by Category

		Frequency	Percent	Cumulative Percent
Valid	Non-Political	2143	46.3	46.3
	Political	2483	53.7	100.0
	Total	4626	100.0	

Table 3.6: Composition of political and non-political news posts in Lankadeepa Facebook Page from 2011 August to 2018 March

Furthermore, the political news posts are subcategorized in to levels depending on areas listed below.

- Functioning of Government and Statutory Boards
- Concessions imposed on the public
- Consumer price changes
- Detection of fraud and corruption
- Collaborative decision making
- Development work performed
- Detection of illicit intoxicants
- Conduct of politicians
- International relations
- Ability of resolving public issues
- Impose law and order
- Introduction of new reforms, regulations and policies
- Protests and strike

A news post can be interacted through reactions, comments or shares. Reactions contains the six animated emojis like, love, wow, sad, angry and haha. Moreover, the reactions are categorized into three groups such as positive, negative and hilarious reactions based on the nature of the emojis.

Positive Reactions – The combination of like, love and wow

Negative Reactions – The combination of sad and angry

Hilarious Reactions – Includes only haha

Total Reactions contains a combination of all six emojis.

The user involvement is measured by the total counts of reactions, comments and shares. The total reactions for each type of news category is shown in Figure 3.18. Although the political news items are higher in count as in Table 3.6, the non-political news items has the highest number of reactions which is twice the value of political news items. This indicates that the users have less interest in political news items compared to other aspects.

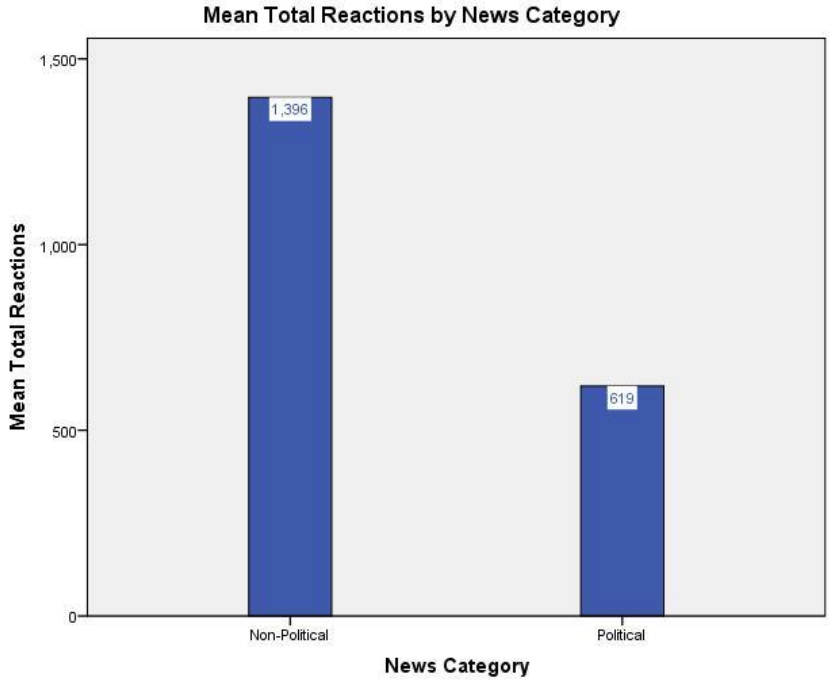

Figure 3.18: Mean Total Interactions including counts of reactions, comments and shares by News Category

Furthermore this can be examined by considering the individual counts for reactions, comments and shares for each category as in Figure 3.19. In all three methods, the non-political news items are higher in count than the political news items which means that user interaction on politics has a downtrend. Contemplating the gap between the counts of political and non-political groups, the highest gap is seen in the shares category and the lowest gap is seen in the comments category which says that, users have shared more news posts of non-political category compared to news posts of political category and they have shown a similar interest in commenting on news items belonging to both categories. On the whole it is visible that, using reactions on a news post is more popular than sharing or commenting on it.

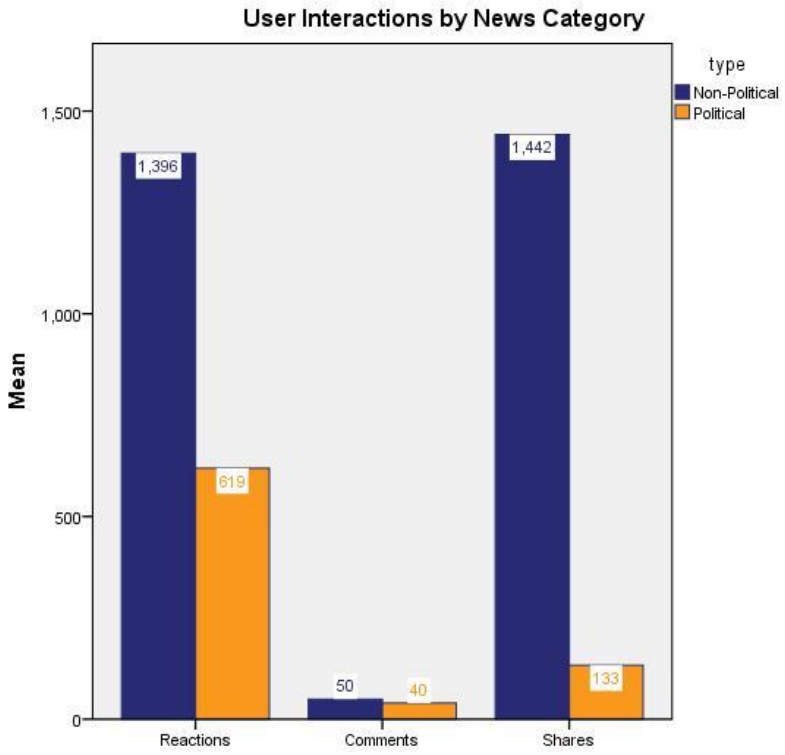

Figure 3.19: Mean User Interactions by News Category

Moreover, the two news categories can grouped by its reaction types from 2016 to 2018 as in Figure 3.20. The like count was not considered as it has a higher value comparatively to other reaction types therefore, results in visual representation issues. In political news items the largest proportion is allotted for “haha” whereas in non-political news items the largest proportion is allotted for “sad”. Similarly, the least proportion for political news items is allotted for “wow” whereas for non-political items the least proportion is allotted for “angry”. Consequently, this shows that the users consider politics as hilarious and less admirable whereas they consider non-politics as empathetic and less aggressive. Majority of non-political news items reported are of social and religious context and this may be the reason for the results driven from the graph.

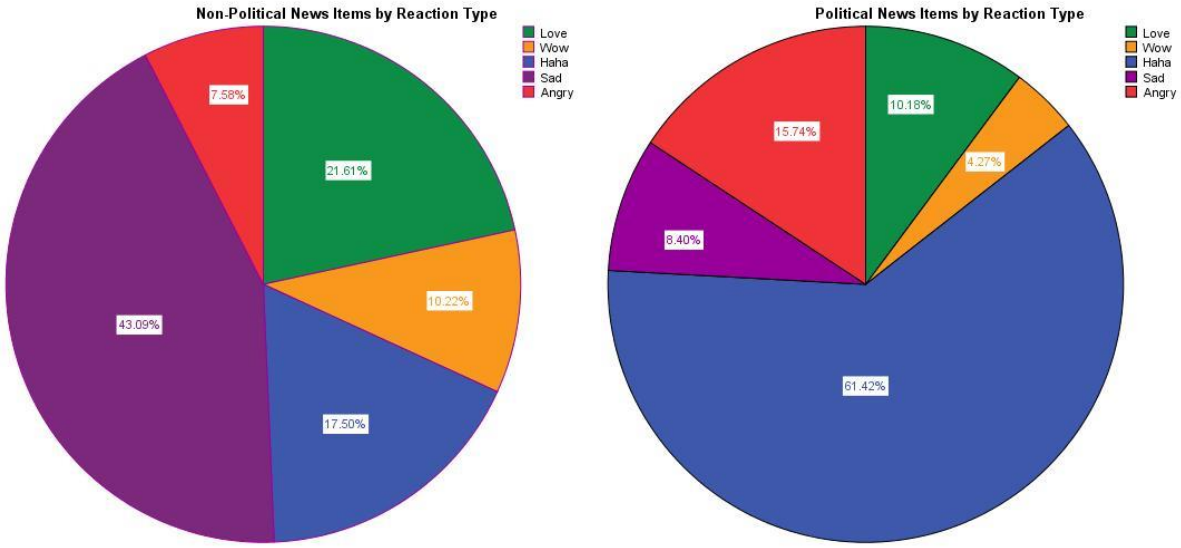

Figure 3.20: Reactions by Type and News Category

As the above results, it is clearly indicated that users are ardent on non-politics over politics. However, following the objectives of this research, the analysis is further narrowed down to find the trends that may exist in the political context.

3.6.3.1 ANALYSIS OF POLITICAL NEWS POSTS

The interactions of political news posts can be plotted against the year as in Figure 3.21 for the aforementioned interaction methods, considering the counts for reactions, comments and shares. All three methods show an uptrend until 2014 followed by a downtrend afterwards. The extremely higher values shown in 2014 is due to the presidential elections held in 2015. Interactions from 2011 to 2014 has higher values compared to interactions from 2015 to 2018. Therefore, it is inevitable that the present government has a lower attentiveness on the contrary of the past government through all three methods. Considering the total comments count by year, a salient pattern cannot be identified between the past or present governance which postulate that the right to free speech is secured by the present governance or breached by the past governance. Hence, it can be assumed that the count of comments are lower than reactions merely due to its ease of use.

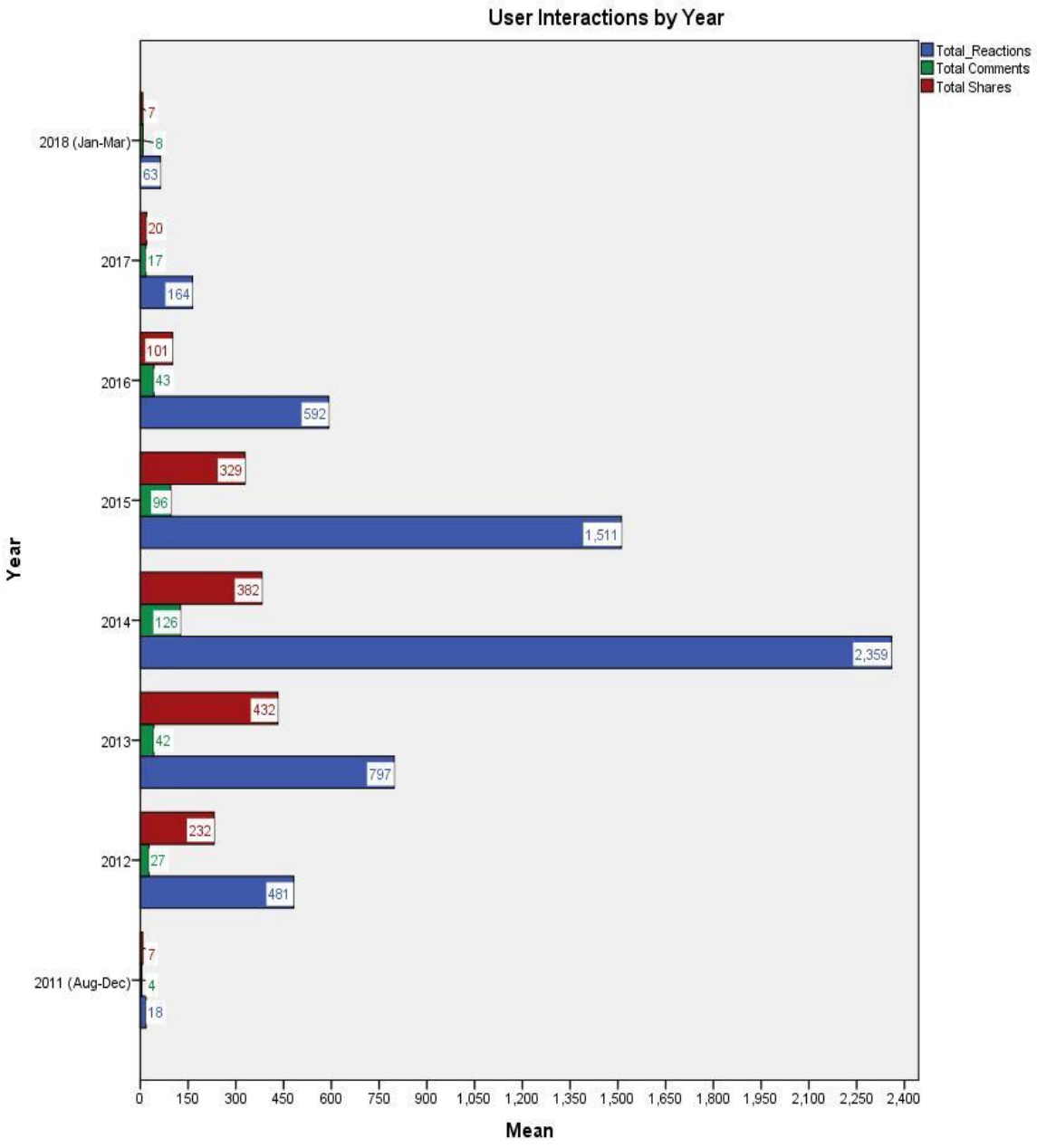

Figure 3.21: Mean User Interactions on Political News Posts by Year

As 2014 shows a peak value in reactions it is of great interest to clarify further the mean total reactions obtained in 2014 by its monthly spread as in Figure 3.22. It shows an increasing pattern over the months from January to December, however the month of May shows an extremely high reaction count compared to other months which is caused by violent, illegal and unjust news reported at that time. It is highly arguable to decide the reason for users to have placed a like on violent news reported. Yet, it may be due to the fact that the only reaction existed at that time was “like”. Additionally, the month of December has a higher value and this is caused by the news reported regarding the Presidential Elections held in 2015. Given these facts, the reaction count does not precisely state that the users have a positive or negative

essence of satisfaction. Thus, it demonstrates the strength of user involvement in political news depending on the state of affairs at the specific time.

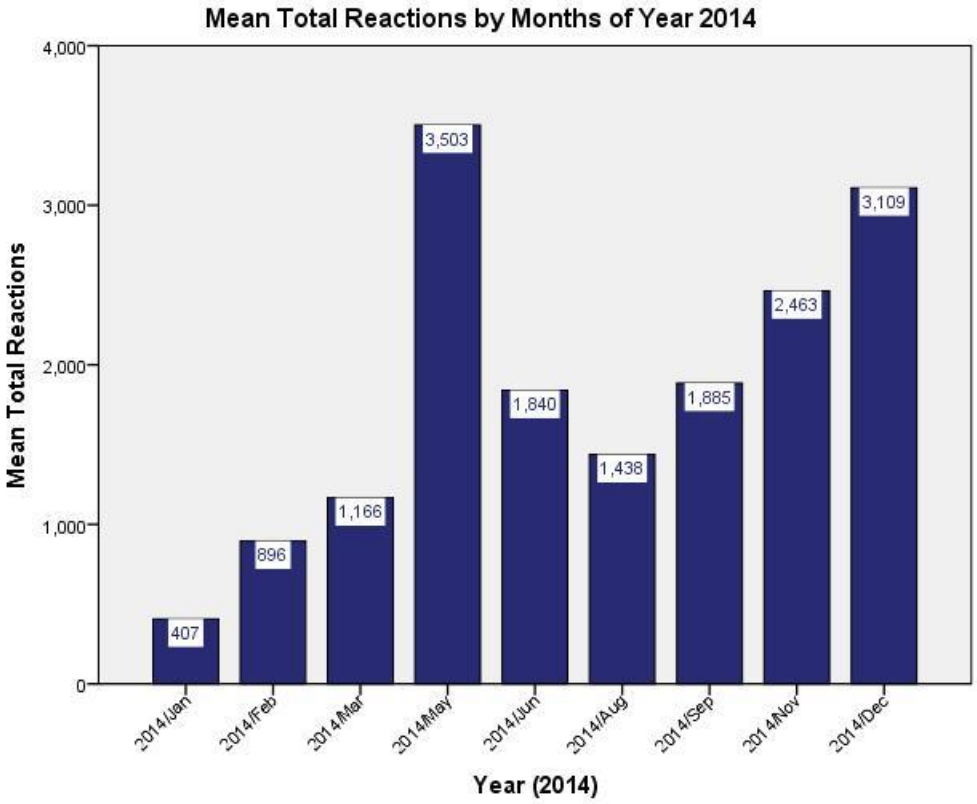

Figure 3.22: Mean Total Reactions by Month for the Year 2014

As Facebook introduced different emojis for reactions in 2016, it is impossible to identify the polarity of reactions before 2016. In Figure 3.23, the stacked bar graph illustrates the positive, negative and total reactions by each year. However, the hilarious reactions are not present precisely. Thus, it is included in the total reactions category. It is seen that the positive reactions decrease and the negative reactions increase from 2016 onwards which indicates that the present government has not met the satisfaction levels of the citizens as expected.

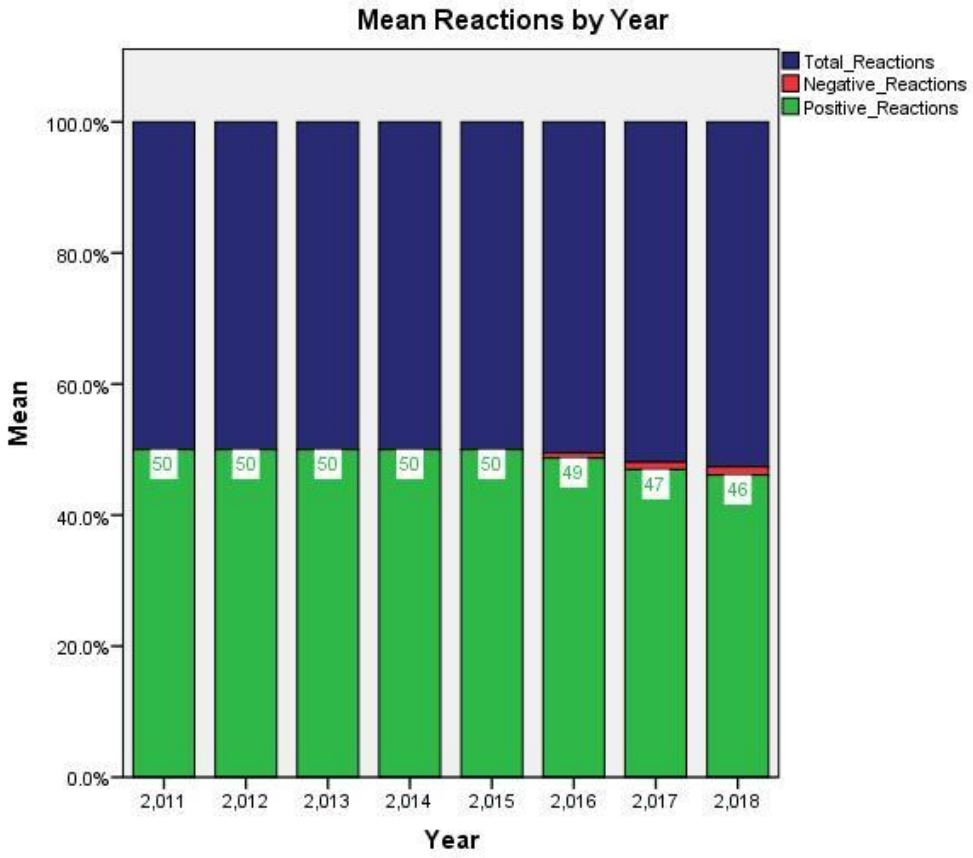

Figure 3.23: Mean value of Total Reaction, Negative Reactions, and Positive Reactions by Year

In order to identify the tendency of reactions for the present government, the mean count for each reaction type is plotted from 2016 to 2018 as in Figure 3.24. As the higher values of the like count is an impediment to represent with the other reactions it was excluded from representation. A higher value for the reaction type “haha” and a lower value for reaction type “wow” is seen compared to other types. It shows that, the majority of users consider politics as hilarious and with least admiration as noted in the section above.

The most number of reactions are identified in the year 2016, which can be due to the enthusiasm and curiosity of the users in exploring reactions as a novel feature. On the whole, all the reactions show a decreasing pattern from 2016 to 2018. Consequently, there can be two reasons, either the citizens are frustrated on the present government and they did not want to interact or they had extreme trust and they hardly wanted to criticize on the government.

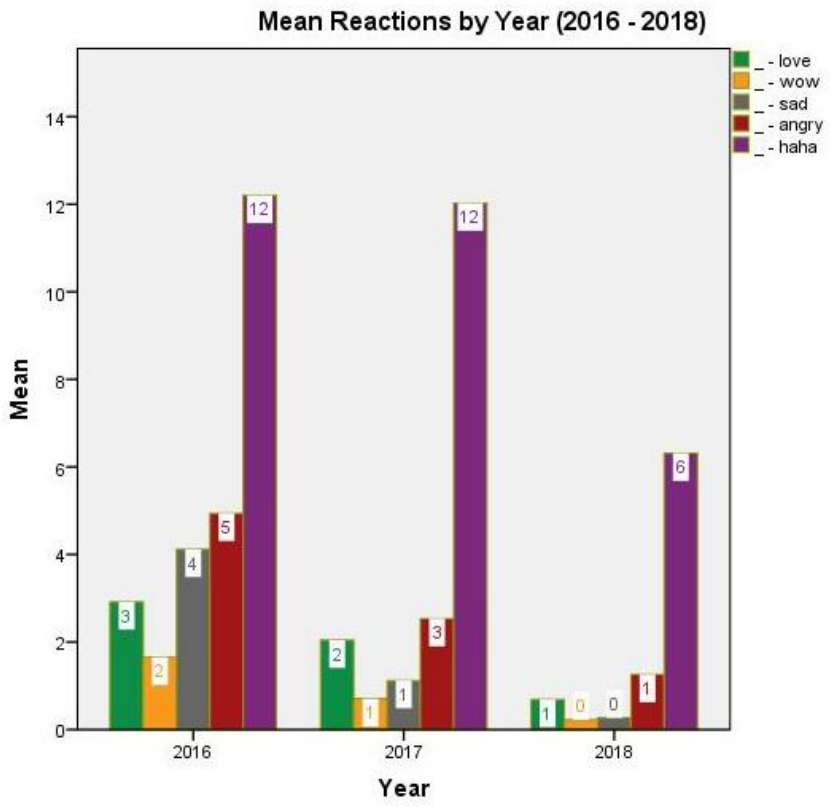

Figure 3.24: Mean Reactions by Type and Year

Moreover, the reactions are visually represented by its positive, negative and hilarious nature as showed in Figure 3.25 for the year 2016, 2017 and 2018. The proportion of positive reactions has a decrement from 2016 to 2018 whereas the proportion of negative reactions has a slight increment and the proportion of hilarious reactions has a substantial increment. Further, it strengthens the claim mentioned above, that the users consider politics as hilarious.

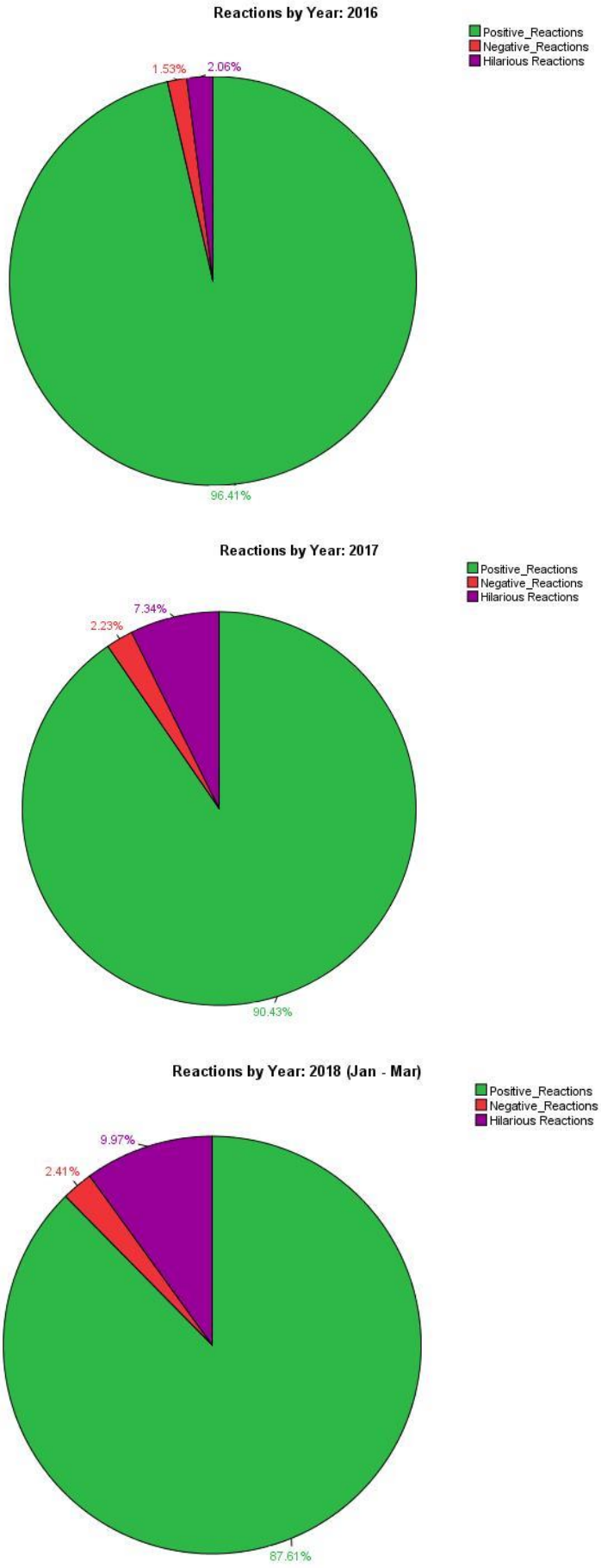

Figure 3.25: Comparison of the Reaction Polarity from Year 2016 to 2018

Reactions can be further analyzed in the presence of a major political event occurred in both governments. The presidential elections held in 2015 and the local government elections held in 2018 has been considered with data of one year ahead of the occurrence of the events. For ease of comparison the timeline can be divided into two segments by a dashed line as in Figure 3.26 in order to differentiate the events. Consequently, the first segment will contain data from January to December in 2014 and the second segment will contain data from January to December in 2017. The first segment has extremely higher number of reactions compared to the second segment and it has an increasing pattern whereas the second segment has extremely lower values and it does not show any increasing pattern, rather it is in a constant state. This may prevail due to many reasons.

- The significance of the type of election, where people were so enthusiastic on the presidential elections over the local government elections.
- The frustration created on the minds of people due to the governance of the present government that they barely react on political posts.
- The satisfaction of people with the governance of the present government with increased trust and seldom expect a change.

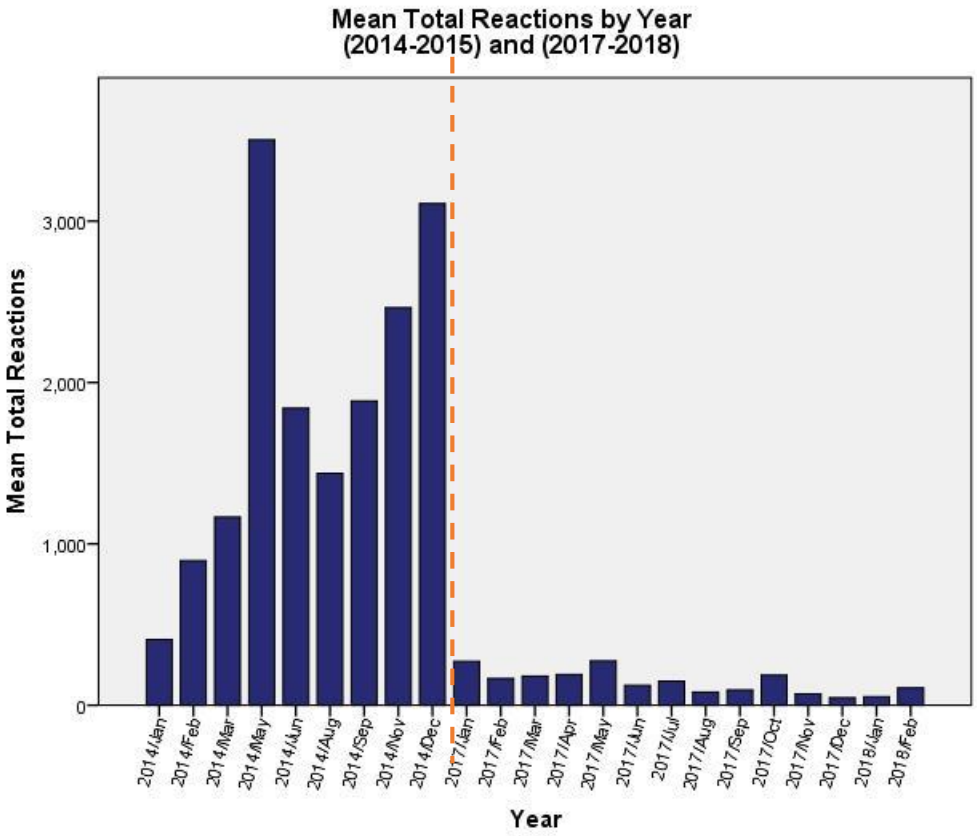

Figure 3.26: Mean Reactions by Year – Comparison of two Political Events

3.6.3.2 ANALYSIS ON THE GOVERNANCE OF THE PRESENT AND PREVIOUS GOVERNMENTS

The present and the past governments can be compared based on their functioning areas. The positive reactions has been used for this since the negative reactions were not present at the reign of the past government. In Figure 3.27 it is seen that, all sub levels has a decreasing pattern in reactions over the ruling years of present government. In contrary, Figure 3.28 shows an increasing pattern in reactions over the ruling years of the past government. In present government, providing concessions, involvement of international relations, resolving public issues and protests have the highest positive reactions in 2015 which means that there have been a prevailing need in fixation of issues in these areas where the previous government failed to attend to.

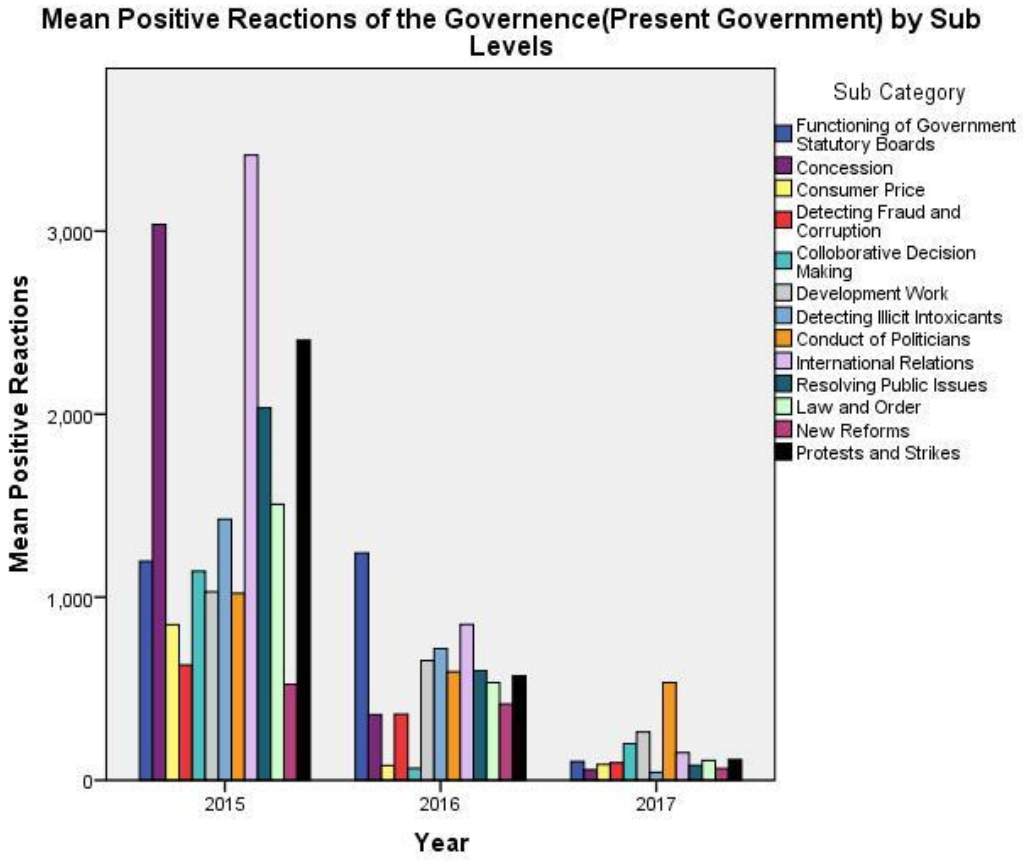

Figure 3.27: Mean Positive Reactions of Present Government by Sub Levels

The past government shows an increment pattern of positive reactions in many areas such as functioning of government and statutory boards, international relations, providing concessions, controlling of consumer prices, detecting illicit intoxicants and protests throughout their governance as seen in Figure 3.28. A sudden raise in positive reactions is observed from 2013

to 2014 which can be due to the presidential election campaigns conducted by the past government in order to gain higher voting levels from the citizens.

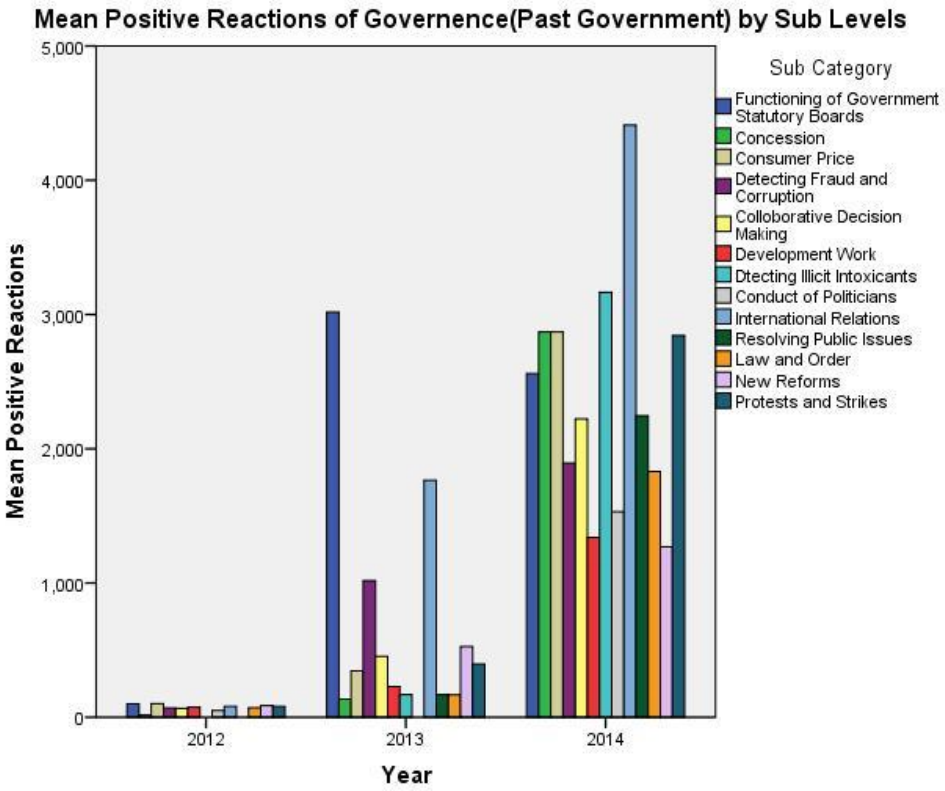

Figure 3.28: Mean Positive Reactions of Past Government by Sub Levels

Similar to the positive reactions, the negative reactions of the present government can be plotted by its sub levels as illustrated in Figure 3.29. In 2016, resolving public issues, detecting illicit intoxicants, detecting fraud and corruption has been in utmost dissatisfaction whereas in 2017, collaborative decision making and conduct of politicians has been in higher dissatisfaction among users. However, it is visible that the areas with negative effects has a decrease from 2016 to 2017 which can be assumed that the present government makes a considerable effort in resolving the prevailing issues.

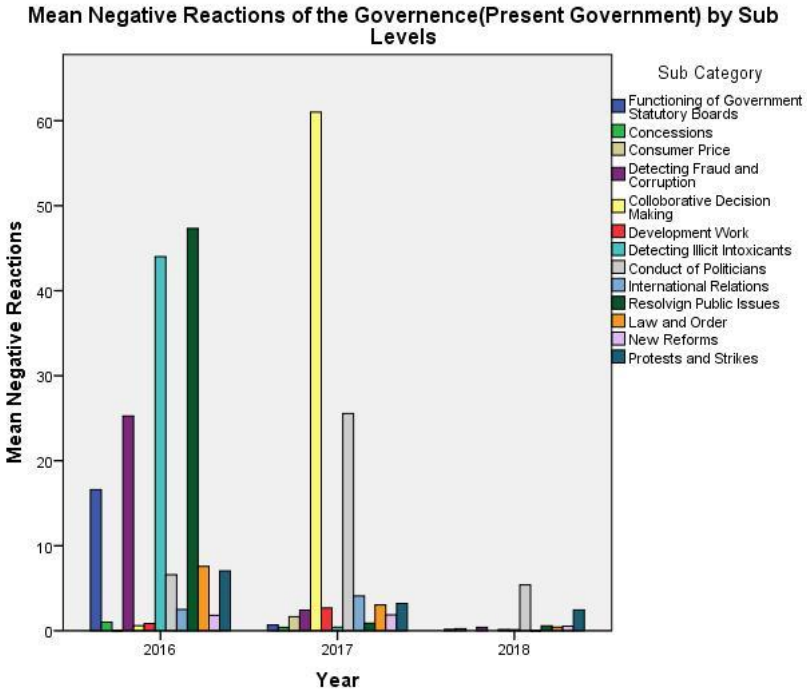

Figure 3.29: Mean Negative Reactions of Present Government by Sub Levels

In comparison to the negative reactions, the hilarious reactions show a substantial increase which can be seen in Figure 3.30. The highest values are observed in 2016 in the sub levels functioning of the government statutory boards, detection of illicit intoxicants and international relations. However, the conduct of politicians remains a constant over the years. The overall hilarious reactions shows a decreasing pattern through 2016 to 2018.

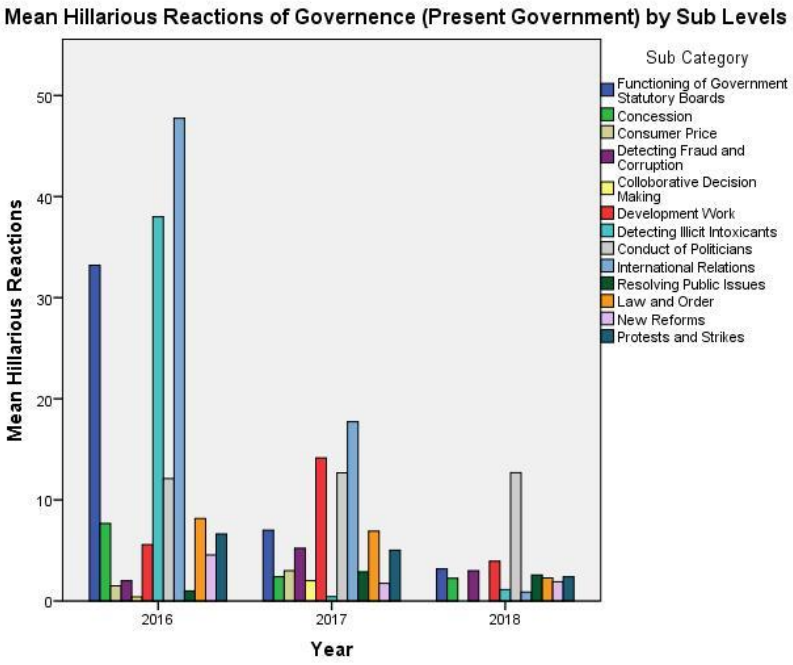

Figure 3.30: Mean Hilarious Reactions of Present Government by Sub Levels

The analysis summarizes the fact that the previous government has an increase interest by the citizens over the present government. By the introduction of the feature of reactions, the present government shows as increase of dissatisfaction by the citizens. However, there is no clear evidence that the higher reaction counts for the past government means that the users were highly appreciating its functioning since by close examination of the news reported at the time of the past government it shows that users have liked posts with violence, unjust and illegal nature. If the reaction emojis were existing at the time of the past government the analysis would have given even more accurate results. The only fact that could be identified is that the high values of the reaction count at the time of the change of governments in 2015 drastically decreases in the period of ruling of the present government, which is not an acceptable standard if the present government claims to have a "Good-Governance". As mentioned throughout this section repeatedly, this can also be due to the high trust level of the citizens that they barely react on the news posts and least expect for a government change. However, such a major decrement cannot be expected from a government which was once highly appreciated by its citizens.

CHAPTER 4 – EVALUATION

This chapter contains the evaluation for this research. As the research suggests to observe the trends of Sri Lankan politics, it is of utmost important to evaluate whether there is any trend in the present government compared to the past government. To achieve this, Mann Kendall test for trend analysis is used. Mann-Kendall Trend Test is a non-parametric test which is used to analyze data collected over time for consistently increasing or decreasing trends. The null hypothesis H0 for these tests is that, there is no trend in the series. The three alternative hypotheses are that, there is a negative, non-null, or positive trend. The test are conducted using the XLSTAT which is an add-on for MS Excel. Data should be prepared in order to apply these tests. Hence, only the positive reactions has been considered with its corresponding month and year values as shown in Table A.6.

The evaluation is done to observe the trend in three cases. All tests are performed using a significance level of 5%.

Case 1: The Overall Trends in Politics

Here, the positive reactions from August, 2011 to March, 2018 has been considered as the data set. In Table 4.1, the summary statistics of the data is shown.

Variable	Observations	Obs. with missing data	Obs. without missing data	Minimum	Maximum	Mean	Std. deviation
14	70	0	70	50.000	295363.000	21677.643	41309.419

Table 4.1: Summary Statistics for Data from Aug, 2011 – Mar, 2018

The hypothesis that is been tested here is whether there is a trend in politics for the entire range. The null and alternative hypothesis can be stated as below.

H0: There is no trend in the series

Ha: There is a trend in the series

The results of the Mann-Kendall Two-tailed Test is illustrated in Table 4.2. As the computed p-value is lower than the significance level alpha=0.05, the null hypothesis H0 can be rejected and the alternative hypothesis Ha can be accepted.

The risk to reject the null hypothesis H0 while it is true is lower than 2.60%.

The Kendall's Tau is the measure of the ordinal association between two measured quantities. The Sen's Slope measures the slope value of a linear trend which in this test is a positive value and holds a magnitude of 85.808.

Kendall's tau	0.182
S	439.000
Var(S)	38908.333
p-value (Two-tailed)	0.026
alpha	0.05
Sen's Slope	85.808

Table 4.2: Mann-Kendall Two-tailed Trend Test

The trend analysis graph for this case is shown in Figure 4.1. The x-axis shows the series values which represents the months of each year.

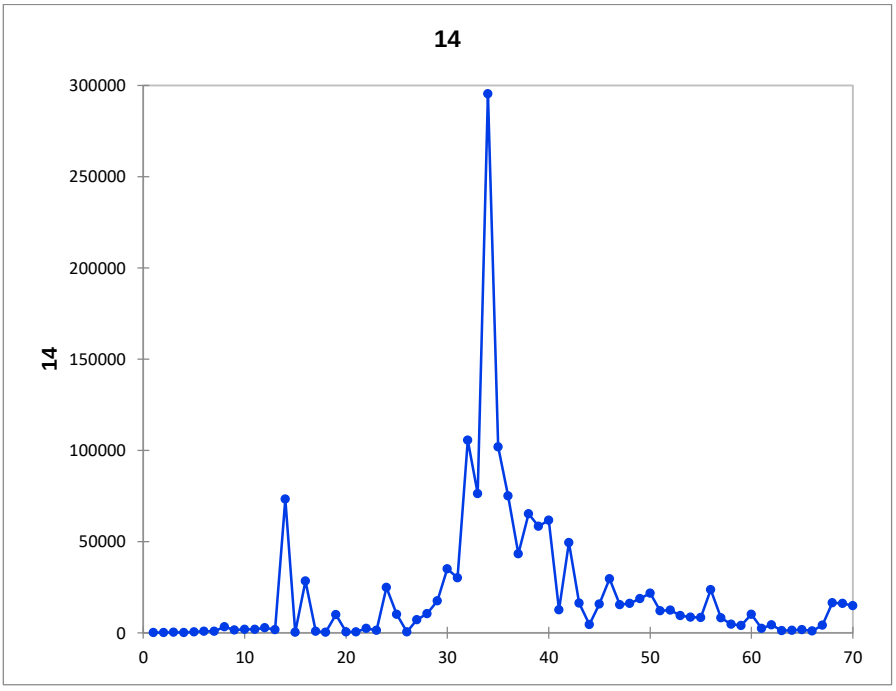

Figure 4.1: Trend Analysis Graph for Positive Reactions from Aug, 2011 – Mar, 2018

Case 2: Trends in Politics for the Past Government

For this case, the positive reactions from 2011, August to 2014, Dec has been considered as the data set. Table 4.3 shows the summary statistics of this test.

Variable	Observations	Obs. with missing data	Obs. without missing data	Minimum	Maximum	Mean	Std. deviation
14	34	0	34	50.000	295363.000	21964.088	54346.521

Table 4.3: Summary Statistics for Data from Aug, 2011 – Dec, 2014

The hypothesis that is been tested here is whether there is a positive trend in politics for the past government. The null and alternative hypothesis can be stated as below.

H0: There is no trend in the series

Ha: There is a positive trend in the series

The Mann-Kendall Upper-tailed Trend test is performed and the results are illustrated in Table 4.4. As the computed p-value is lower than the significance level alpha=0.05, the null hypothesis H0 is rejected and the alternative hypothesis Ha is accepted.

The risk to reject the null hypothesis H0 while it is true is lower than 0.01%.

The Sen's slope holds a positive value with a magnitude of 488.5. It is a much higher value compared to the value identified in case 1.

Kendall's tau	0.558
S	313.000
Var(S)	4550.333
p-value (one-tailed)	< 0.0001
alpha	0.05
Sen's Slope	488.500

Table 4.4: Mann-Kendall Upper-tailed Trend test

The trend analysis graph is shown in Figure 4.2 which has a monotonically increasing trend. The x-axis shows the data values which represents the months of each year from Aug, 2011 to Dec, 2014.

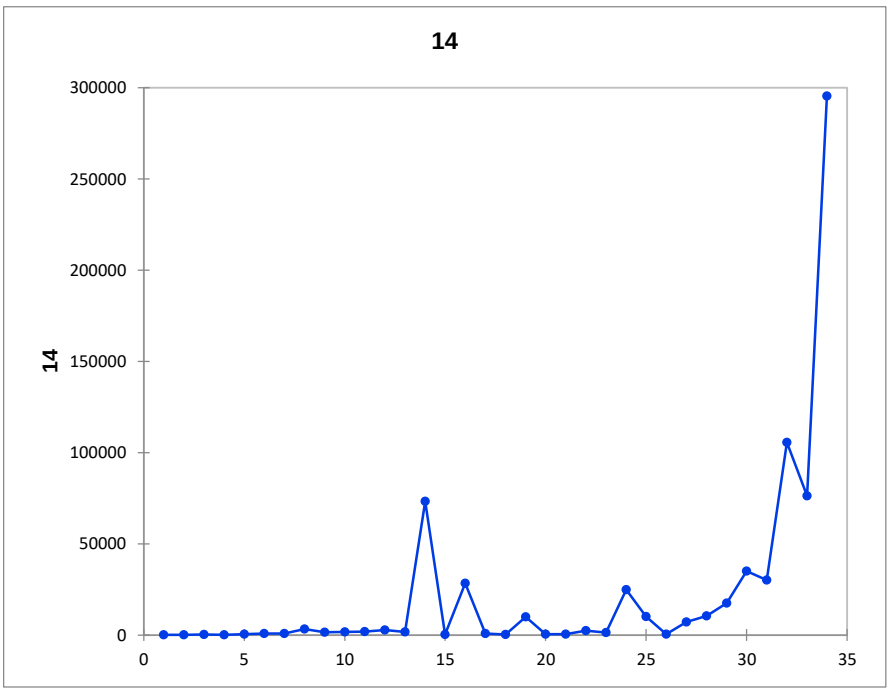

Figure 4.2: Trend Analysis Graph for Positive Reactions from Aug, 2011 – Dec, 2014

Case 3: Trend in Politics for the Present Government

In this case, the positive reactions from January, 2015 to March, 2018 has been considered as the dataset. In Table 4.5, the summary statistics of the data is shown.

Variable	Observations	Obs. with missing data	Obs. without missing data	Minimum	Maximum	Mean	Std. deviation
101852	35	0	35	948.000	74999.000	19108.686	20012.959

Table 4.5: Summary Statistics for Data from Jan, 2015 – Mar, 2018

The hypothesis that is been tested here is whether there is a negative trend in politics for the ruling period of the present government. The null and alternative hypothesis can be stated as below.

H0: There is no trend in the series

Ha: There is a negative trend in the series

The Mann-Kendall Lower-tailed Trend test performed is illustrated in Table 4.6. As the computed p-value is lower than the significance level alpha=0.05, the null hypothesis H0 is rejected and the alternative hypothesis Ha is accepted.

The risk to reject the null hypothesis H0 while it is true is lower than 0.01%.

The Sen's slope holds a negative value with a magnitude of 874.

Kendall's tau	-0.556
S	-331.000
Var(S)	4958.333
p-value (one-tailed)	< 0.0001
alpha	0.05
Sen's Slope	-874.000

Table 4.6: Mann-Kendall Lower-tailed Trend test

The trend analysis graph for this case is shown in Figure 4.3 which has a monotonically decreasing trend. The x-axis shows the data values which represents the months of each year from Jan, 2015 to Mar, 2018.

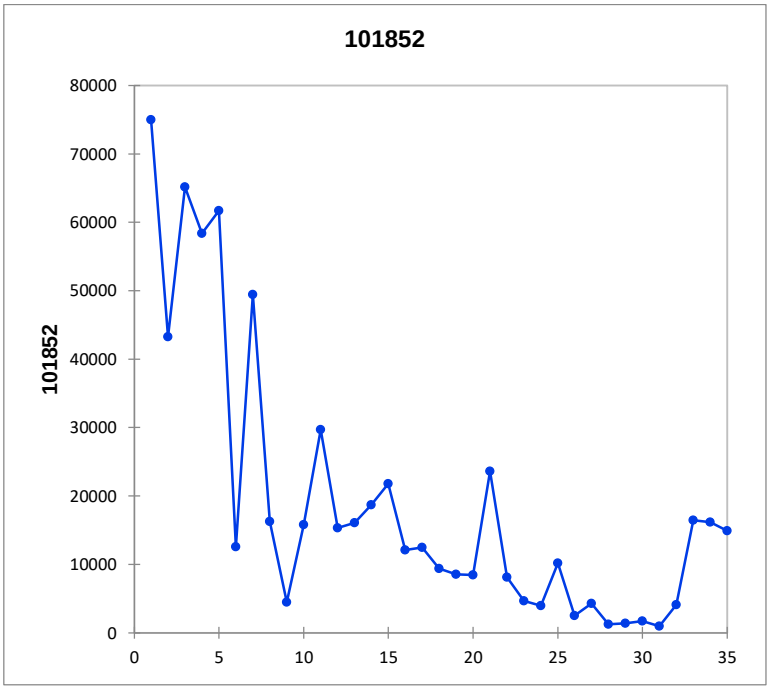

Figure 4.3: Trend Analysis Graph for Positive Reactions from Jan, 2015 – Mar, 2018

In conclusion, it is seen that on the whole political news reported has an increasing trend in positive reactions over the years from 2011 to 2018. However, if the time periods of past and present governments are considered separately, it can be stated that the past government has a monotonically increasing trend whereas the present government has a monotonically

decreasing trend. The Sen's Slope values for the trends further strengths the visually represented trends for the two governments which indicates that the past government has a high positive slope value compared to the present government having a high negative slope value. The evaluation proves the results driven from the analysis done for political news posts in Chapter 3.

CHAPTER 5 – CONCLUSION AND FUTURE WORK

In this thesis the main objective was to identify any trends in Sri Lankan politics in the perspective of its citizens. Facebook, which is the most popular social media platform at present was chosen to retrieve data. Through observation, Lankadeepa was identified as the most popular news page used by many active Sri Lankan online users. The data of interest for this research is the news posts and their reaction, comment and share counts. Additionally, the creation time of the posts were also obtained. Four thousand six hundred twenty six records were obtained for the analysis adhering to certain Facebook limitations. The data were pre-processed to remove inconsequential posts which includes cover photo and profile photo updates and unescaped characters. The news posts are grouped depending on their nature as political and non-political. Furthermore, the political news posts are subcategorized in to sub levels considering areas on functioning of government and statutory boards, concessions imposed on the public, consumer price changes, detection of fraud and corruption, collaborative decision making, development work performed, detection of illicit intoxicants, conduct of politicians, international relations, ability of resolving public issues, impose law and order, introduction of new reforms, regulations and policies, protests and strike. The news posts collected includes data of past and present governments and is further segmented to obtain interesting conclusions.

As the news items are categorized based on my personal opinion inevitably a bias is imposed and a survey was conducted to find the bias. Thus, a sample was chosen considering the religion and gender. Personal details such as age and education level were collected from each respondent in addition to the religion and gender. However, Hindus were not considered for this analysis due to their illiteracy in Sinhala which in fact is a requisite to read the news posts. Three types of survey forms were created comprising of hundred news items gathered from Lankadeepa news posts which each type included all political, all non-political and a mixture of political and non-political news posts where the respondents were asked to classify them into their appropriate categories. The bias is calculated by acquiring the total responses classified as political, against my categorization for political news. Subsequently, the overall total bias calculated was twenty eight. Additionally, an analysis was performed regarding the survey to find the factors that affected the bias imposed. The results showed a lower divergence in bias where the respondents were males of age group 31-35 holding a Bachelors' degree and above. This concludes that, the educational background, age and gender are prime factors affecting the reasoning ability of a person. Religion was not a vital factor however, various norms imposed

by certain religious cultures may have an impact for acquiring further education. Hence, indirectly religion becomes a considerable factor.

The analysis performed regarding the news posts includes identifying trends of the user involvement in the context of politics and non-politics, comparison of the user reactions on past and present government bodies. The results show that, users are less interested in politics and consider it as hilarious. It was also observed that in general, the user reactions on politics has a monotonically increasing trend throughout the entire period whereas the present government has a monotonically decreasing and past government has a monotonically increasing trend.

In this research, only the reaction counts were considered to generate the trends in political news items. As future work, this can be extended to read the content of the comments using a sentimental analysis and get the polarity for each post and validate it with the trends obtained using reactions. This will be useful to measure the accuracy and reliability of the trends in both methods. In addition to the trends generated for political news items, trends can be generated for non-political news items as well, depending on its social and religious aspects. Moreover, the political news posts can be subcategorized into further levels in various dimensions based on its nature which can be useful to perform a more rigorous analysis.

REFERENCES

[1] C. Bucur, "Opinion Mining Platform for Intelligence in Business," *conomic Insights – Trends and Challenges Vol.III (LXVI) No. 3/2014,* pp. 99 - 108, 2014.

[2] P. H. D. Saurin Dave, "Trend Analysis in Social Networking using Opinion Mining A Survey," *IJSRSET,* vol. 1, no. 6, 2015.

[3] G. P. A. G. Michail Salampasis, "Using social media for continuous monitoring and mining of consumer behaviour," *International Journal of Electronic Business,* vol. 11, no. 1, 2014.

[4] D. T. D. ProfessorChristopher G.Reddicka, "A social media text analytics framework for double-loop learning for citizen-centric public services: A case study of a local government Facebook use," *Government Information Quarterly,* vol. 34, no. 1, pp. 110-125, January 2017.

[5] M. M. G. F. S. Mariam Adedoyin-Olowe, "A Survey of Data Mining Techniques for Social Media Analysis," arxiv.org, Aberdeen, Reading, 2013.

[6] A.-E. M. S. S. Mhamdi C., "Text Mining and Analytics: A Case Study from News Channels Posts on Facebook," Springer, Cham, 2018.

[7] K. B. Diego Tumitan, "Tracking Sentiment Evolution on User-Generated Content:A Case Study on the Brazilian Political Scene," Brazil, 2013.

[8] K. L. H. H. Sethunya R Joseph, "Social Media Data Mining: An Analysis & Overview of Social Media Networks," *International Journal of Database Theory and Application,* vol. Vol.9, no. 7, pp. 291-296, 2016.

[9] A. M. L. &. M. B. O. Julia K. Woolley, "The 2008 Presidential Election, 2.0: A Content Analysis of User-Generated Political Facebook Groups," *Mass Communication and Society ,* vol. 13, no. 5, pp. 631-652, 2010.

[10] R. C. P. Juliet E. Carlisle, "Is Social Media Changing How We Understand Political Engagement? An Analysis of Facebook and the 2008 Presidential Election," *Political Research Quarterly,* vol. 66, no. 4, pp. 883-895 , 2013.

[11] L. D.-X. Stefan Stieglitz, "Social media and political communication: a social media analytics framework," *Social Network Analysis and Mining,* vol. 3, no. 4, p. 1277–1291, 2013.

[12] L. C. S. M. I. G. P. Andrea Ceron, "Every tweet counts? How sentiment analysis of social media can improve our knowledge of citizens' political preferences with an application to Italy and France," *SAGE Journals,* vol. 16, no. 2, pp. 340-358, March 1, 2014.

[13] S. H. Dekay, "How large companies react to negative Facebook comments," *Corporate Communications: An International Journal,* vol. 17, no. 3, pp. 289-299, 2012.

[14] T. A. Small, "WHAT THE HASHTAG? A content analysis of Canadian politics on Twitter," *Networking Democracy?: Social media innovations and participatory politics*, vol. 14, no. 6, pp. 872-895, 2011.

[15] ""Census of Population and Housing of Sri Lanka, 2012 - Table A4: Population by district, religion and sex," Department of Census & Statistics, Sri Lanka, 2012. [Online]. Available: http://www.statistics.gov.lk/PopHouSat/CPH2011/Pages/Activities/Reports/FinalReport/Population/Table%20A4.pdf. [Accessed 6 February 2018].

[16] "Sri Lanka Census of Population and Housing," Department of Census and Statistics, Sri Lanka, 2012. [Online]. Available: http://sis.statistics.gov.lk/statHtml/statHtml.do?orgId=144&tblId=DT_POP_SER_266&conn_path=I2. [Accessed 6 February 2018].

APPENDIX A

This appendix consists of the data definitions of the data sets obtained through Facebook and the survey forms. Table A.1 is a high level categorization of a news item whether it is political or non-political.

Category	Symbol
Political	1
Non-Political	0

Table A.1: News items categorization as politics and non-politics

Table A.2, A.3, A.4, A.5, A.6 are the data definitions and the notations assigned for the fields collected in the survey form.

Category	Symbol
Male	M
Female	F

Table A.2: Categories of Gender in the survey form

Category	Symbol
Buddhism	B
Christianity	C
Islam	M
Aethiesm/ Agnostism/ Univerlism	A

Table A.3: Categories of Religion in the survey form

Category	Symbol
Advanced Level	A
Diploma	D
Higher Diploma	H
Bachelor’s Degree	B
Post Graduate Diploma	P
Master’s Degree	M
Master of Philosophy	F
Professional Qualifications	O

Table A.4: Categories of Highest Educational Qualification in the survey form

Category	Symbol
Yes	Y
No	N

Table A.5: Categories of Following External Degree/Diploma/Professional Qualifications Apart from the Highest Educational Qualifications in the survey form

Row Labels	Sum of Positive_Reactions
2011/Aug	14
2011/Sep	50
2011/Oct	132
2011/Nov	229
2012/Jan	130
2012/Feb	495
2012/Mar	863
2012/Apr	752
2012/May	3376
2012/Jun	1574
2012/Jul	1800
2012/Aug	1812
2012/Sep	2681
2012/Oct	1759
2012/Nov	73339
2013/Jan	322
2013/Feb	28346
2013/Mar	809
2013/Apr	261
2013/May	10025
2013/Jun	510
2013/Jul	521
2013/Aug	2436
2013/Sep	1397
2013/Nov	24909
2013/Dec	10201
2014/Jan	407
2014/Feb	7164
2014/Mar	10497
2014/May	17517
2014/Jun	34968
2014/Aug	30189
2014/Sep	105578
2014/Nov	76367
2014/Dec	295363
2015/Feb	101852
2015/Mar	74999
2015/May	43270
2015/Jun	65148
2015/Aug	58392
2015/Sep	61709
2015/Oct	12547
2015/Nov	49449
2015/Dec	16239

2016/Jan	4481
2016/Feb	15804
2016/Mar	29683
2016/Apr	15344
2016/May	16098
2016/Jun	18690
2016/Jul	21794
2016/Aug	12098
2016/Sep	12462
2016/Oct	9395
2016/Nov	8521
2016/Dec	8444
2017/Jan	23625
2017/Feb	8131
2017/Mar	4652
2017/Apr	3943
2017/May	10198
2017/Jun	2496
2017/Jul	4277
2017/Aug	1253
2017/Sep	1371
2017/Oct	1723
2017/Nov	948
2017/Dec	4082
2018/Jan	16470
2018/Feb	16155
2018/Mar	14913
Grand Total	**1517449**

Table A.6: Pre-processed data with the sum of positive reactions by its month and year

APPENDIX B

This appendix describes data retrieval from Facebook using the supplementary source code.

Figure B.1, illustrates the source code used, to get data from the Facebook Page "Lankadeepa". PHP SDK for Facebook has been used to integrate Facebook login and make requests to the Graph API (version 2.11). Prior to making API requests, a Facebook application should be created and obtain the App ID and App Secret. Using the Access Token Tool, a User Token can be obtained. By using the URL of the Facebook Page "Lankadeepa" in the Graph API Explorer, the id of the page could be obtained which is denoted by '`181270011941164`'. These parameters are used to send the request to the API including the required query. The response is iterated to form the JSON format needed. The `offset` and `limit` are variables that should be manually changed after each execution of the script.

```
<?php
require_once __DIR__ . '\vendor\facebook\graph-
sdk\src\Facebook\autoload.php';
use Facebook\FacebookRequest;
use Facebook\Facebook;
use Facebook\FacebookApp;
use Facebook\GraphNodes\GraphNodeFactory;

$fb = new Facebook([
  'app_id' => '2038255509738021',
  'app_secret' => 'd1ec6e378987e3ef097099345e5efdda',
  'default_graph_version' => 'v2.11',
  ]);

  $fbApp = new FacebookApp($fb->getApp()->getId(), $fb->getApp()-
>getSecret());

  $request = new FacebookRequest($fbApp,
'EAAc9yFeekiUBAC0e5FVZBbWUFZBSJF6flZCZBi0tjAwcUiMYNm19SW9lSKY1Nnht6dWvNairl
WZBly6tlYxZCZBOBsNyXEa6SfTFAR74pAE3pfUDLvfk8EVWj3pMzt1JSa0zTdRwqpMk28eZBLta
daPMBzPhLt9KDmlJMRM5rmcBNRcDaagJ3a99GIHkS90Lz9oZD', 'GET',
'181270011941164/feed?fields=id,created_time,shares,comments.limit(0).summa
ry(total_count),name,message,reactions.type(LIKE).summary(total_count).limi
t(0).as(like),reactions.type(LOVE).summary(total_count).limit(0).as(love),r
eactions.type(WOW).summary(total_count).limit(0).as(wow),reactions.type(HAH
A).summary(total_count).limit(0).as(haha),reactions.type(SAD).summary(total
_count).limit(0).as(sad),reactions.type(ANGRY).summary(total_count).limit(0
).as(angry)&limit=100&offset=0');

try {
  $response = $fb->getClient()->sendRequest($request);

} catch(Facebook\Exceptions\FacebookResponseException $e) {
  echo 'Graph returned an error: ' . $e->getMessage();
  exit;
} catch(Facebook\Exceptions\FacebookSDKException $e) {
  echo 'Facebook SDK returned an error: ' . $e->getMessage();
  exit;
}
```

```
$i=0;
$results = $response->getDecodedBody() ;

while($i < 100) {

    echo '{"id":"'.$results["data"][$i]["id"].'",';

    if(isset($results["data"][$i]["created_time"])){
        echo '"created_time":"'.$results["data"][$i]["created_time"].'",';
    }

    if(isset($results["data"][$i]["message"])){
        echo '"title":"'.$results["data"][$i]["message"].'",';
    }else if(isset($results["data"][$i]["name"])){
        echo '"title":"'.$results["data"][$i]["name"].'",';
    }else{
        echo '"title": "NIL",';
    }

    if(isset($results["data"][$i]["like"])){
        echo '"like":'.$results["data"][$i]["like"]["summary"]
["total_count"].",";
    }else{
        echo '"like": 0,';
    }

    if(isset($results["data"][$i]["love"])){
        echo '"love":'.$results["data"][$i]["love"]["summary"]
["total_count"].",";
    }else{
        echo '"love": 0,';
    }

    if(isset($results["data"][$i]["haha"])){
        echo '"haha":'.$results["data"][$i]["haha"]["summary"]
["total_count"].",";
    }else{
        echo '"haha": 0,';
    }

    if(isset($results["data"][$i]["sad"])){
        echo '"sad":'.$results["data"][$i]["sad"]["summary"]
["total_count"].",";
    }else{
        echo '"sad": 0,';
    }

    if(isset($results["data"][$i]["wow"])){
        echo '"wow":'.$results["data"][$i]["wow"]["summary"]
["total_count"].",";
    }else{
        echo '"wow": 0,';
    }

    if(isset($results["data"][$i]["angry"])){
        echo '"angry":'.$results["data"][$i]["angry"]["summary"]
["total_count"].",";
    }else{
        echo '"angry": 0,';
    }

    if(isset($results["data"][$i]["shares"]["count"])){
        echo '"shares":'.$results["data"][$i]["shares"]["count"].",";
```

```
    }else{
        echo '"shares": 0,';
    }
    echo '"comments":'.$results["data"][$i]["comments"]["summary"]
["total_count"];
    echo "},<br>";
    $i++;
}
 ?>
```

Figure B.1: Source Code for Data Retrieval from Facebook Page "Lankadeepa"